NOAA Technical Report NWS 55

Decision Support Focused on Improving Water Quality: Development of the Runoff Risk Advisory Forecast for Wisconsin Farmers

Dustin C. Goering

North Central River Forecast Center

February 2014

U.S. DEPARTMENT OF COMMERCE

National Oceanic and Atmospheric Administration

National Weather Service

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AA	Analysis Accumulation
AFO	Animal Feeding Operation
ARS	Agricultural Research Service
ASOS	Automated Surface Observing System
AWOS	Automated Weather Observing System
BFI	Base Flow Index
BMP	Best Management Practice
CAFO	Concentrated Animal Feeding Operation
CAT 1	Category 1 or block of time when no runoff event simulated
CAT 2	Category 2 or simulated events below basin threshold
CAT 3	Category 3 or simulated events at or above basin threshold
CNRFC	California-Nevada River Forecast Center
COOP	Cooperative Observer Program
CSI	Critical Success Index
DATCP	Department of Agriculture, Trade, and Consumer Protection
DNR	Department of Natural Resources
DOA	Department of Agriculture
DOC	Department of Commerce
EOF	Edge-of-Field
EPA	Environmental Protection Agency
ESP	Ensemble Streamflow Prediction
ET	Evapotranspiration
F0	Simulated runoff event generated from rainfall only
F1	Simulated runoff event generated from rainfall and snowmelt
F2	Simulated runoff event generated from snowmelt only
FAR	False Alarm Ratio
FFG	Flash Flood Guidance
FMAP	Future Mean Areal Precipitation
FMAT	Future Mean Areal Temperature
HAS	Hydrologic Analysis and Support Forecaster at a RFC
HEC-RAS	Hydraulic Model from U.S. Army Corps of Engineers
HL-RDHM	Hydrology Lab Research Distributed Hydrologic Model
INTRO	Interflow Runoff component from SAC-SMA
LDM	Liquid Dairy Manure
MAP	Mean Areal Precipitation

MAT	Mean Areal Temperature
MMTF	Manure Management Task Force
NAEFS	North American Ensemble Forecast System
NCRFC	North Central River Forecast Center
NDFD	NWS National Digital Forecast Database
NGOM	Northern Gulf of Mexico
NMP	Nutrient Management Plan
NOAA	National Oceanic and Atmospheric Administration
NOHRSC	National Operational Hydrologic Remote Sensing Center
NRCS	Natural Resources Conservation Service
NWS	National Weather Service
NWSLI	National Weather Service Location Identifier
OE	Observed Runoff Event
OHD	Office of Hydrologic Development
POD	Probability of Detection
QPF	Quantitative Precipitation Forecast
RAIM	Rain + Melt component from Snow-17
RFC	River Forecast Center
RRAF	Runoff Risk Advisory Forecast
SAC-HTET	SAC-SMA with enhanced Evapotranspiration and Heat Transfer
SAC-SMA	Sacramento Soil Moisture and Accounting Model
SBM	Solid Beef Manure
SE	Simulated Runoff Event
USDA	United States Department of Agriculture
USGS	United States Geological Survey
UZFW	Upper Zone Free Water from SAC-SMA
UZK	UZFW depletion coefficient from SAC-SMA
UZTW	Upper Zone Tension Water from SAC-SMA
UZTWC	Upper Zone Tension Water Contents from SAC-SMA
UZTWD	Upper Zone Tension Water Deficit from SAC-SMA
UZTWM	Upper Zone Tension Water Maximum from SAC-SMA
WD	Warning Day
WDNR	Wisconsin Department of Natural Resources
WFO	Weather Forecast Office
WPC	Weather Prediction Center
WPDES	Wisconsin Pollutant Discharge Elimination System
WRN	Weather Ready Nation

ABSTRACT

The Runoff Risk Advisory Forecast (RRAF) provides Wisconsin's farmers with an innovative decision support tool which communicates the threat of undesirable conditions for manure and nutrient spreading for up to 10 days in advance. The RRAF is a pioneering example of applying the National Weather Service's hydrologic forecasting abilities towards the Nation's water quality challenges. Relying on the North Central River Forecast Center's (NCRFC) operational Snow17 and Sacramento Soil Moisture Accounting Models, runoff risk is predicted for 216 modeled watersheds in Wisconsin. The RRAF is the first-of-its-kind real-time forecast tool to incorporate 5-days of future precipitation as well as 10-days of forecast temperatures to generate runoff risk guidance. The forecast product is updated three times daily and hosted on the Wisconsin Department of Agriculture, Trade, and Consumer Protection (DATCP) website. Developed with inter-agency collaboration, the RRAF model was validated against both edge-of-field observed runoff as well as small USGS gauged basin response. This analysis indicated promising results with a Bias Score of 0.93 and a False Alarm Ratio (FAR) of only 0.34 after applying a threshold method. Although the threshold process did dampen the Probability of Detection (POD) from 0.71 to 0.53, it was found that the magnitude of the events categorized as hits was 10-times larger than those classified as misses. The encouraging results from this first generation tool are

aiding State of Wisconsin officials in increasing awareness of risky runoff conditions to help minimize contaminated agriculture runoff from entering the State's water bodies.

1. INTRODUCTION

1.1 Project Concept

Runoff from agricultural fields has long been known to carry excessive amounts of nutrients and sediments into nearby streams, rivers, lakes and groundwater reservoirs. This common problem is often compounded across many parts of the country when farmers spread livestock manure on fields (USEPA, 1990). Improper spreading techniques, inadequate or no nutrient management plans, and unfortunate timing can often lead to freshly spread manure being carried off fields, contaminating nearby water bodies, and eventually enhancing the hypoxia zone in the Gulf of Mexico (USEPA, 2002).

There is a wealth of information available to manure producers and spreaders to help plan for and manage their livestock waste (from this point the terms "producers" and "spreaders" will refer to anyone spreading manure, whether it is farmers or commercial manure spreading operations). Best management practices (BMP), peer experience, and guidance from local, state and, federal agencies are available for producers to incorporate into a management plan. However, guidance relating to another significant component in managing manure applications is often missing in the spreader's toolkit. That major component is the future runoff risk inherent in their fields.

It is assumed many responsible spreaders rely on themselves to collect information regarding their current soil conditions as well as expected weather conditions. They must use that information to formulate their decision on when and where to spread manure. Simply walking their fields can help them inventory the conditions of their fields. Weather information can be found many places these days such as television, radio stations, or websites.

However, gathering the mentioned information and relying on themselves to process it only provides the spreader with a qualitative assessment. This assessment could be incomplete or not representative of the expected atmosphere-soil system. The suggested decision making process detailed above, probably the real-world best case scenario, still allows the possibility of poor decisions to be made regarding spreading manure. Possible poor decisions could result in dangerous levels of nutrients transported from fields directly into the aquatic system. These incidents carry negative impacts both to the spreader in terms of financial penalties (Good, 2012) as well as to the environment in terms of quality degradation leading to aquatic life kills in rivers, lakes, and the ocean (USEPA, 2002).

This project's goal is to provide spreaders with a forward looking decision support tool that will incorporate the complex interaction of future temperatures, precipitation, and soil conditions resulting in a real-time runoff risk assessment for their area. Complementing their existing management guidelines, the Runoff Risk

Advisory Forecast (RRAF) will help to minimize the occurrence of contaminated runoff events thus maximizing the benefit of spreading manure across agriculture fields. The RRAF accomplishes this by addressing the largest unknown in managing manure spreading: when and when not to spread because of weather conditions.

This study documents the development and implementation of the first-of-its-kind RRAF across Wisconsin. The long term success of the product will depend on two questions. Will the RRAF be an accurate predictor of average field scale conditions and runoff risk across a given basin? Will spreaders buy into the product and build trust in the guidance thus resulting in fewer incidents of contaminated runoff?

1.2 Project Motivation

The development of the RRAF stems from events dating back to the winter of 2004 to 2005. That winter Wisconsin witnessed a large number of manure runoff events across the state that generated a lot of public interest. In fact, during the period from July 1st, 2004 through June 30th, 2005, there were 52 runoff events where manure derived impacts were documented. A vast majority of them, 62%, occurred in February and March (12 and 20 events respectively). Spreading manure over fields was found to be the cause of 74% of these events. Frozen ground or snow covered conditions were present for 84% of these incidents, while another 8% occurred during saturated soil conditions or rain. These contaminated runoff events had the following damages associated with them: fish kills in 17%, well contamination in 20%, and discharge to water bodies in 43%. Associated livestock operations covered the spectrum of sizes, however most were under the Concentrated Animal Feeding Operation (CAFOs) size and therefore not directly regulated by state agencies at that time (WDNR & DATCP, 2006). CAFOs will be discussed some more in Chapter 2.

Following that damaging spring the Secretaries of Wisconsin's Department of Natural Resources (DNR) and Department of Agriculture, Trade, and Consumer Protection (DATCP) formed a Manure Management Task Force (MMTF) in May of 2005 made up of 16 members from diverse backgrounds. The Task Force was asked to provide a report offering advice on reducing risks of manure runoff incidents with emphasis on reducing the acute runoff events from land applied manure that damage surface and ground water. In the desire to remain neutral and not punitive, the Task Force was also instructed to consider recommendations that provided a balance between protecting the environment as well as allowing a favorable climate for the important livestock industry to grow and prosper in the state (WDNR & DATCP, 2006).

By definition, the MMTF described acute events as those that deliver large amounts of nutrients and toxins from manure to water sources within hours of application. Common conditions that were found to contribute to these acute incidents included spreading manure on frozen or snow covered ground, spreading on saturated soils, and spreading immediately before rain or snowmelt events (WDNR & DATCP, 2006).

The final report submitted by the MMTF in March of 2006 proposed 8 recommendations for the state agencies and livestock industry to implement in order to reduce manure contaminated runoff events. Of particular significance to the development of what was to become the RRAF were parts (b) and (c) of recommendation 4 listed below:

> *b. Developing a manure spreading advisory system that may take the form of a web-based risk assessment tool to warn farmers about specific weather related hazards such as predicted rain events.*
>
> *c. Developing a statewide notification program to alert farmers concerning high risk spreading conditions such as melt periods and dry weather. Different media including radio broadcasts (e.g., daily market reports), websites, and email could be used for making notifications.*

At that time, DATCP had searched around the country and could not find any examples of a risk assessment tool that was described in 4.b. There were a couple of examples based almost entirely on future precipitation and described more in section 2.6. The National Weather Service (NWS) entered into the picture in early 2008 during an outreach event where hydrologists from the North Central River Forecast Center (NCRFC) were presenting information on their services to interested people in Wisconsin. A member of the audience noted the combination of the soil and runoff models with forecast elements and connected the NCRFC with representatives in DATCP. Shortly after, a working group associated with the tasking from the MMTF report, led by DATCP, formed to begin investigating the possibility of the NCRFC providing daily forecast data to support the RRAF.

1.3 Applying National Weather Service Capabilities

The National Weather Service, an agency under the National Oceanic and Atmospheric Administration (NOAA), is the Federal agency that provides official weather forecast and climate services. Not as commonly known in the public arena is the fact that the NWS is also responsible for providing forecasts for lake levels as well as river streamflow and stages across the nation. The alignment of NOAA with the Department of Commerce (DOC) supplies a second component to the NWS mission: along with protecting life and property, the NWS also focuses on aiding the national economy. The NWS accomplishes the components of its mission in the hydrology sphere via the work at thirteen River Forecast Centers (RFCs) across the United States.

As the NWS developed its strategy for the 21st century, documented in the Weather Ready Nation (WRN) strategic plan (NWS, 2011), it was determined the agency needs to shift its focus away from being strictly a supplier of weather and water products. The new focus will center on assisting decision makers at various levels in specific arenas where weather and water forecasts have an effect. Essentially getting to know the customer, their problems, and how NWS services can help them make better decisions is the new business model.

The RRAF is a unique and exciting example of a decision support tool that meshes with several of the 8 goals listed in the WRN road map to help fulfill the new NWS vision (NWS, 2011). It is one of the first forecast tools aimed at the increasingly important area of water quality. Historically the NWS, through the RFCs, have concentrated on providing streamflow and point location stage forecasts and derived from these, river flood watches and warnings.

The second exciting development is that the NWS is introducing itself to an entirely new customer base that would not have looked to the NWS for guidance before. By entering into the realm of water quality and providing forecast services that land managers can use and apply, the NWS becomes even more relevant and useful to the entire livestock management community.

A noteworthy achievement of the RRAF development and implementation was demonstrating the ability to leverage existing RFC modeling capabilities to provide new services and products. This was a major test to the NWS ability to move in this direction as limited financial resources and heavy operational demands suggested no new resources would be available for a product like the RRAF to be created at this time.

1.4 Description of North Central River Forecast Center (NCRFC)

The North Central River Forecast Center is one of thirteen RFCs covering the United States. The NCRFC area of responsibility spans nine states, 341,357 mi^2, and three major watersheds. The first drainage area is the Hudson Bay watershed which includes Devils Lake, the Souris River, and the Red River of the North. The second is the drainages around the western Great Lakes. The third is the Upper Mississippi River watershed down to Chester, IL. The office is staffed by 19 people with 14 of them involved with daily forecast operations and responsible for monitoring 426 forecast point locations (Figure 1). Three of the 14 operational positions are designated as Hydrologic Analysis and Support (HAS) forecasters which are primarily involved with meteorological data processing (observed and forecast) as well as some river forecasting. The remaining 11 positions are hydrologic forecasters with primary duties of maintaining and operating the hydrologic modeling in the office. The NCRFC is open every day at a minimum from 0600L to 2200L. During major flooding events the office transitions to a 24-hour schedule.

The NCRFC runs many types of models across its area of responsibility. The hydrologic model currently used is the Sacramento Soil Moisture Accounting model (SAC-SMA) which has been in continuous operation at the RFC for nearly twenty years. Another model used is the Snow-17 model. This model handles the precipitation typing and snowpack simulation producing a rain + melt (RAIM) time series that is entered into the SAC-SMA. The SAC-SMA and Snow-17 are the two models used in the production of the RRAF. Other models in use at the NCRFC include routing routines such as Tatum to distribute streamflow to downstream locations, a unit hydrograph model, and reservoir routines where needed. Many reaches along major rivers have hydraulic models such as the HEC-RAS in use.

The models are set up and ran in a lumped basin approach where basins are generally defined where historical or present river gauge data are available. The basin delineations can change over time as new gauge data becomes available or new forecast services are needed. The current configuration has 1,173 sub-watersheds parsed into 33 larger watersheds referred to as “Forecast Groups”. On a daily basis these Forecast Groups are assigned to available forecasters who are then responsible for reviewing those basins and making adjustments if necessary before issuing required forecasts. As of December 2012, 842 sources of river stage data were being used meaning nearly 72 percent of the basins were gauged. The smallest modeled watershed was 6.5 mi^2 while the largest was 3,061 mi^2. The mean basin size was 291 mi^2 with a standard deviation of 281 mi^2.

Models are run several times per day with a forecaster review on the 12Z, 18Z, and 00Z model runs daily. During these runs forecasters evaluate the model forcing data (observed and forecast precipitation and temperature) as well as the models themselves. Modifications can be made on basin by basin basis to many of the models to align the model states with observed data or regional behavior. The objective of NCRFC forecasters is to accurately simulate streamflow at a particular basin while maintaining reasonable model states. Forecaster adjustments are only desired when they are necessary and justified. In order to provide the most useful stage and flow forecasts possible, it is essential the forecasters maintain accurate model states.

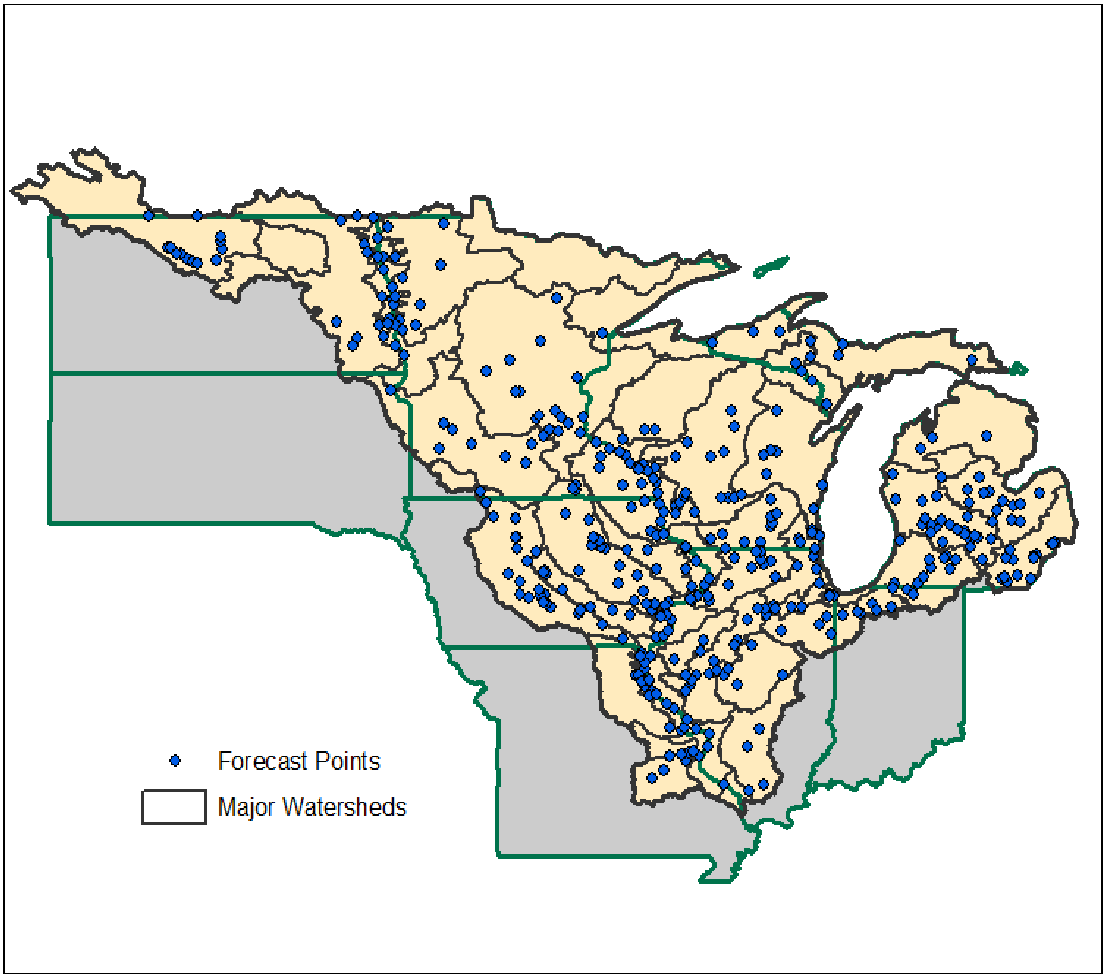


Figure 1. Distribution of the 426 Forecast Points in the NCRFC Region.

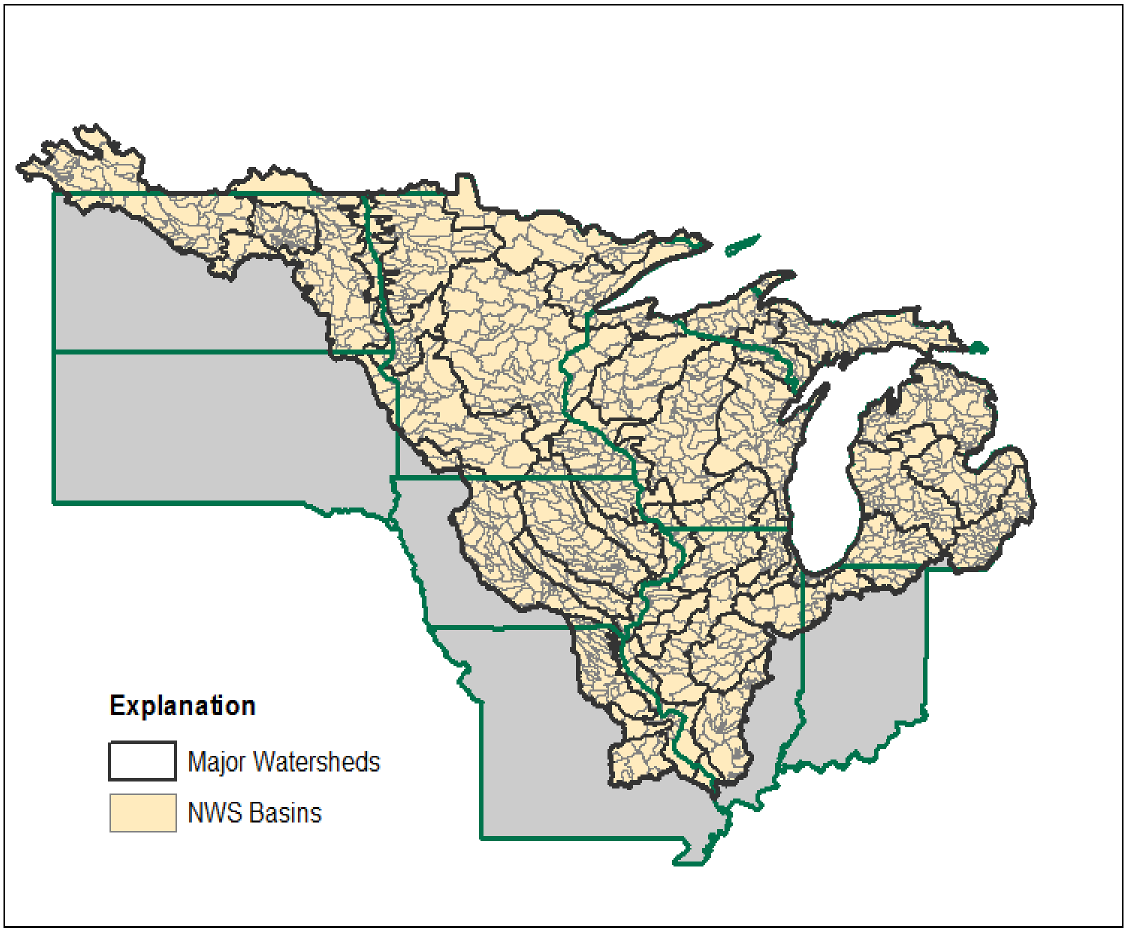


Figure 2. Map of the 1,173 NCRFC modeled sub-watersheds and 33 Forecast Groups spread across nine states.

2. ASPECTS OF MANURE MANAGEMENT

2.1 Wisconsin Dairy Summary

Agriculture and raising livestock have long been a fundamental necessity for societies for millennia and to this day continue to occur in every state in the United States. In 1997 the United States Department of Agriculture (USDA) reported that there were over 1.2 million farms with livestock and poultry operations in the U.S. (USEPA, 2002) (USEPA, 2012 [b]). Focusing only on Wisconsin, the USDA tallied over 77,000 farms covering 15 million acres of farmland as of the end of 2011 (USDA, 2011 [a]). According to the Wisconsin Department of Natural Resources (DNR) livestock totals on these farms include nearly 3.5 million cattle of which 1.265 million are dairy cows. There were over 3.5 million turkeys and over 7 million chickens being raised as well as 340,000 hogs and pigs, and 84,000 sheep as of the end of 2012 (WDNR, 2012 [b]). In total, Wisconsin's agriculture sector employs over 350,000 people and generates $59.6 billion to the state's economy (WDNR, 2012 [a]).

Although the RRAF product has widespread applications for limiting contaminated runoff from fields in Wisconsin, the primary focus is on the state's dairy producers and the application of liquid dairy manure (LDM). Additional background information presented will emphasize production and manure generation from dairy farms in Wisconsin.

The average milk cow produces over 20,000 pounds of milk per year (USDA, 2011 [a]). For Wisconsin as a whole, that puts production over 27 billion pounds annually. In 2007, milk and other dairy products generated over $4.5 billion in sales to rank second in that category in the U. S. (USDA, 2011 [a]). That amount of production was equivalent to 51% of the total agricultural sales for Wisconsin. To provide perspective, sales from various grains in the state only generated $1.6 billion (18% of the annual total), while cattle and calves sales were third with just over $1 billion (11 %) (USDA, 2011 [a]). Figure 3 presents the spatial distribution of dairy cows across Wisconsin by county based on USDA data from the end of 2011 (USDA, 2011 [b]).

The latest trend in livestock production is the consolidation of smaller livestock operations to fewer number of larger herd size operations across the U.S. In 2007 the United States Environmental Protection Agency (EPA) estimated that of the over 1 million livestock farms in the U.S. more than 200,000 of them are considered Animal Feeding Operations (AFOs) (USEPA, 2012 [b]). The designation of a farm being an AFO depends on the animals being kept and raised in confinement for all or part of the year. The EPA reports that since 2003 the number of AFOs nationally has decreased, but the number of animals contained in these operations has increased (USEPA, 2012 [b]) . The Wisconsin DNR and agencies in neighboring Minnesota (UM-Extension, 2012 [b]) have also observed this trend across their states in the last ten years (WDNR, 2012 [b]). In fact, the University of Minnesota Extension has reported that Minnesota has lost 50% of their dairies with 200 cows or less since 1993 and that half of the national dairy herd resides in operations of 500 or more cows while less than 10% reside on farms of 50 animals or less (UM-Extension, 2012 [b]).

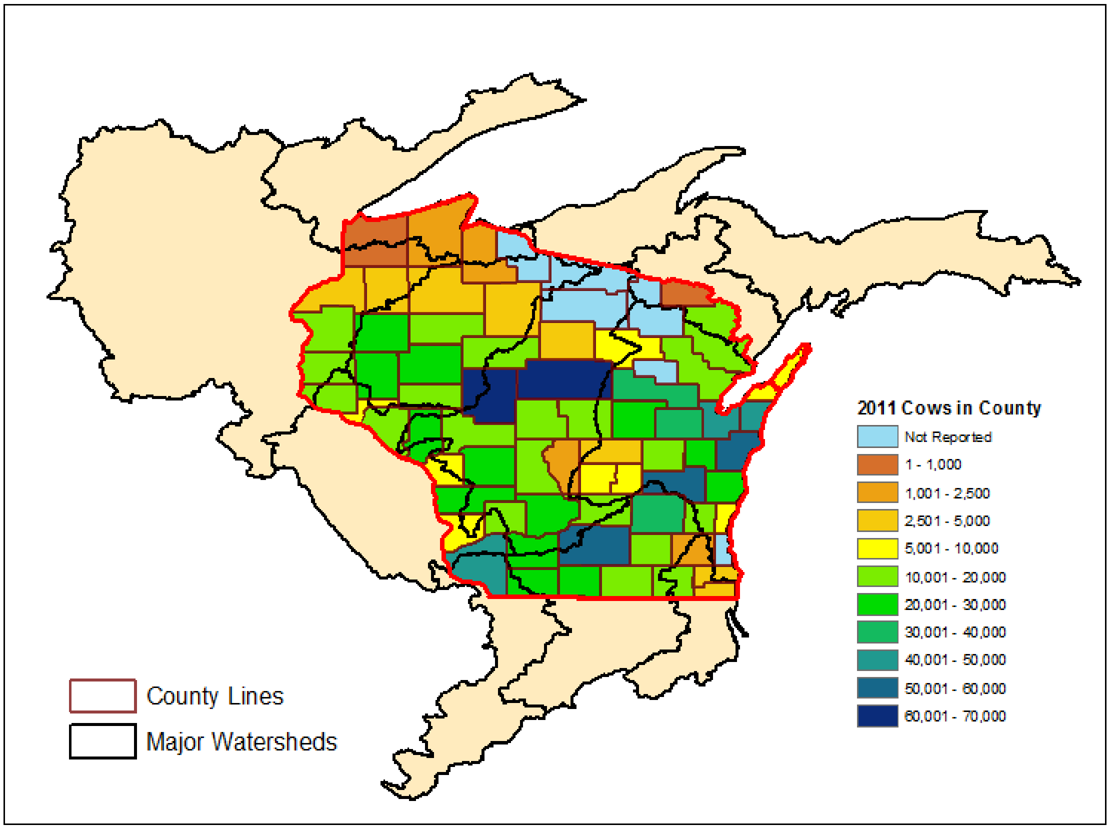


Figure 3. Number of dairy cows per county in 2011. Data provided by USDA National Agricultural Statistics Service (USDA, 2011 [b]).

Even higher density livestock operations are deemed Concentrated Animal Feeding Operations (CAFO). Operations are defined as a "Large CAFO" if the farm has more than 1,000 animal units. One animal unit is equal to one 600 pound steer (WDNR, 2012 [b]). In terms of dairy operations the threshold for a large CAFO in animal units is equivalent to 700 mature dairy cows (USEPA, 2002). The significance of the large CAFO designation is that the state is required to regulate manure storage and application for these larger operations under the U.S. EPA Clean Water Act (WDNR, 2012 [a]). Nationwide the U.S. EPA estimates there were over 15,000 CAFOs in 1997 (USEPA, 2002). In Wisconsin there were 233 CAFOs of which 217 were dairy farms at the end of 2011 (WDNR, 2012 [a]). Figure 4 provided by the Wisconsin DNR for the spatial distribution of these CAFOs across their state (WDNR, 2011).

The result of dairy production across Wisconsin is not only the hefty sales numbers and boost to the economy but also the significant amount of manure generation. For perspective on the scale of waste generated by milk cows it is valuable to first become familiar with their required daily intakes of food and water. A typical lactating Holstein cow consumes between 50 and 80 pounds of feed and requires between 18 and 36 gallons (150 – 300 pounds) of water daily. That same cow in return will provide around 70 to 150 pounds (8 - 17 gallons) of milk per day (UM-Extension, 2012 [b]).

Manure generation obviously is the other byproduct and the sheer amount produced by large dairy herds across the country requires serious attention to negate environmental hazards as well as reap any nutritional benefits it provides. In 1997 the USDA estimated total manure generation from all livestock and poultry in the U.S. totaled over 1 billion tons which is six times more than human-generated waste (USEPA, 2002). In addition the USDA and the EPA have consistently estimated manure generation from AFOs to be near 500 million tons per year (USEPA, 2002) (USEPA, 2012 [b]). Focusing on the individual cow scale, a typical mature dairy cow weighs 1,400 pounds and will produce 148 pounds (17.7 gallons) of waste a day. That adds up to over 54,020 pounds (27 tons) or 6,560 gallons of waste annually per cow (UW-Extension, 2012). As alluded to earlier, dairy cows and milk production lead the way in waste generation. For example a typical 1,100 pound beef steer will produce 80 pounds of waste and a 150 pound hog will produce only 9.5 pounds of waste daily. Annually that is 14.6 tons for one beef steer and only 1.7 tons per hog (UW-Extension, 2012).

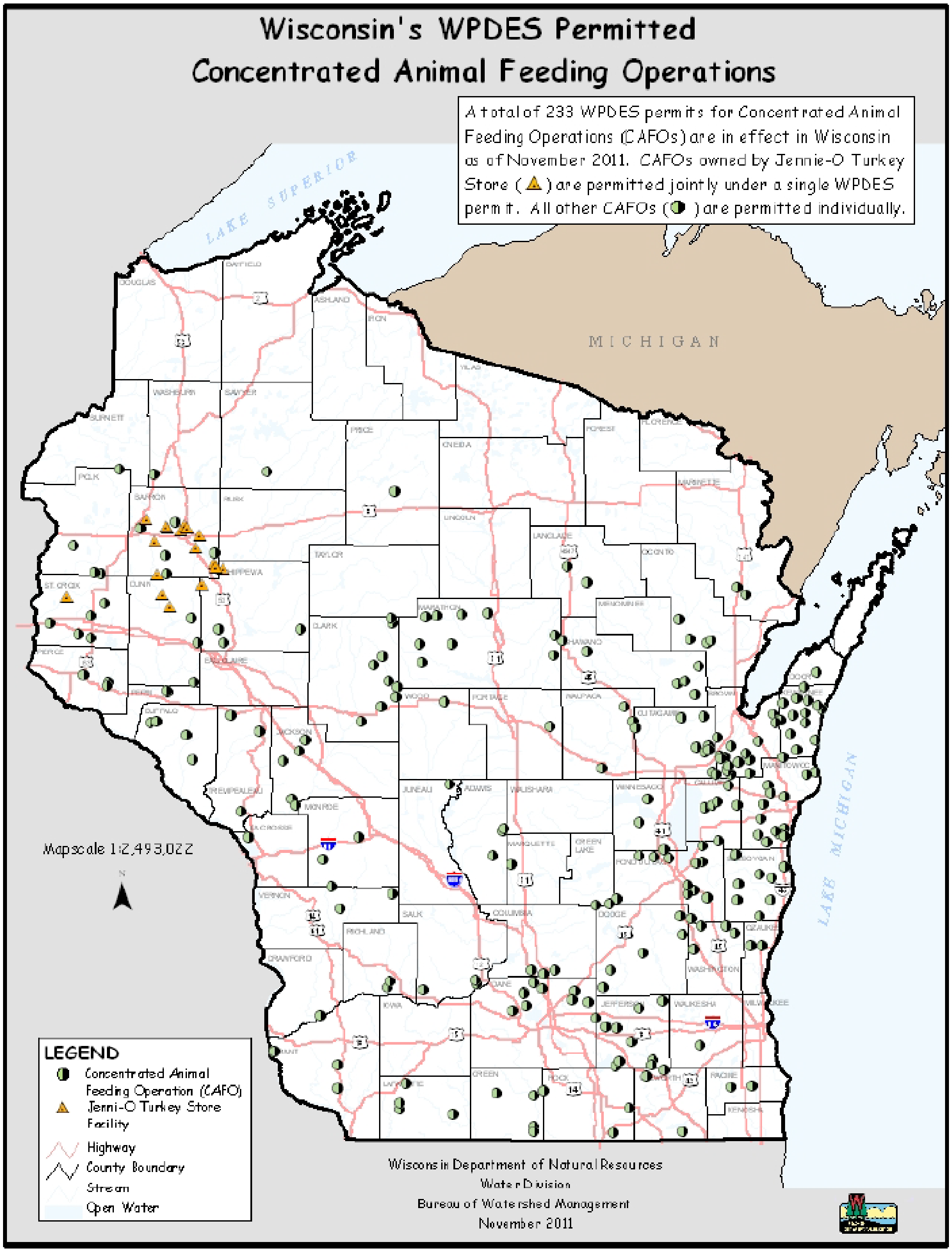


Figure 4. Concentrated Animal Feeding Operations (CAFOs) in Wisconsin as of November 2011 (WDNR, 2011).

Applying these annual waste production values to the Wisconsin dairy herd as a whole suggests that over 8 billion gallons (or over 34 million tons) is generated from milk and dairy production alone. The University of Wisconsin Extension estimates the state dairy industry produces near 12 billion gallons of waste annually. Helping to visualize that amount, they mention it would fill a standard football field to a depth of 5.25 miles deep (UM-Extension, 2010). The difference in computed volumes is theorized to occur due to the lower estimate derived strictly from waste generated from the animal, whereas the larger 12 billion gallons could include the additional waste water that is generated from cleaning various facilities, often referred to as process wastewater. For example milking parlors are generally cleaned and sanitized after each milking. Cleaning of other animal storage and bedding locations creates contaminated water which is often diverted to the manure and waste storage areas.

It should also be noted that not all dairy farmers handle the transportation and field application of their manure. The University of Wisconsin Extension estimates that nearly one third of the 12 billion gallons of manure produced annually is spread by 116 custom (or for hire) applicator businesses (UM-Extension, 2010). Success of the RRAF product will depend on buy-in and routine use not just by the producers themselves but also the custom hauler/applicator industry as well.

2.2 Advantages of Manure Application

2.2.1 Cost Benefits

Although the value of agricultural production in Wisconsin sounds impressive, and undoubtedly has a significant impact on the state's overall economy, it is not hard to imagine that on a farm-by-farm basis the margin between making profit and losing money is becoming thinner by the year. As with any industry, defined requirements and associated costs for daily operations exist in the agriculture sector. Power, fuel, equipment, wages, and fertilizer are just a few of the types of expenses crop and livestock farmers encounter.

Generally, the public has little ability to manage significant costs in their daily lives or operations. However, many farmers differ in this regard when they begin to view their manure waste as an asset and not as an expense. This changing point of view has also been heightened by recent increases in commodity prices. Farmers are beginning to see their crops worth more and beginning to realize they can effectively trim costs by maximizing their manure usage instead of relying on synthetic fertilizers. In the early part of last decade it was not uncommon for dairies to plead with neighbors to take manure off their hands (UW-Extension, 2009). However after around 2009, the rising costs of synthetic fertilizers caused a dramatic shift in the manure market towards the producers' favor. It is now common for producers to sell their manure to the highest bidder as well as crop farmers to actively inquire about and seek manure from livestock owners (UW-

Extension, 2009). This larger scale economic effect of fertilizer prices has turned manure into a source of income for many livestock operations (UM-Extension, 2012 [d]).

The University of Wisconsin Discovery Farms has noted that with current fertilizer prices (all per pound: $0.53 for nitrogen, $0.50 for phosphate, and $0.48 for potash) a herd of 100 cows is producing manure worth $17,000 per year as fertilizer (Discovery Farms, 2012 [b]). Prior to the change in the fertilizer expenses, manure application costs often exceeded the value of the nutrients applied (UM-Extension, 2012 [d]).

Rising costs of fertilizers are resulting in encouraging changes in farmers' attitudes and behaviors with respect to manure application. As manure has become a primary fertilizer, it is in the farmers' financial interest to maximize the economic value and benefits of the resource. Therefore farmers are more motivated to follow proper application guidelines keeping manure on their fields (Discovery Farms, 2011). An example of application guidelines maximizing costs could be the common practice of incorporating manure within three days of application. In economic terms, nitrogen losses from not incorporating can reach 3 pounds per 1,000 gallons of manure. Assuming an application rate of 10,000 gallons per acre, not incorporating within three days could result in a loss of $12 to $24 per acre depending on the price of nitrogen (Dickrell, 2009). Environmental agencies are also beginning to take notice and promote the idea of thinking of manure management in terms of economics (Gauldin, 2008). Doing this reinforces the mutually positive couplet of higher crop yields and lower costs for the farmer as well as environmental benefits.

2.2.2 Soil Benefits

The practice of applying livestock manure to agricultural fields has been something done by farmers for a very long time. Farmers have been aware that manure improves soil quality for centuries (UM-Extension, 2011 [a]). Within the modern era, scientists have begun to understand many of the benefits manure provides to the soil and thus the crops. Essentially manure acts as both a fertilizer and soil conditioner. It provides essential nutrients needed by crops such as nitrogen and phosphorus, but also improves the soil structure.

2.2.2.1 Conditioning the Soil

Manure is by nature a great source of organic matter, which helps improve soil structure and enables it to both support nutrient sources and transport them to plants (UM-Extension, 2011 [a]). This allows the soil to resist compaction and increases its ability to hold water (USEPA, 2012 [b]). Increasing water holding capacity decreases crop stress and soil erosion as well as increasing nutrient retention (UM-Extension, 2011 [a]). Manure is also a good source and stimulant for microbial activity in soils (Farm Journal, 2010). Studies have also shown that manure applications to fields have reduced both soil erosion and runoff (UM-

Extension, 2011 [a]). It has been shown that manure applied to fields can behave similar to crop residue in protecting the soil and limiting soil particle detachment and erosion due to reducing rain drop impacts (UM-Extension, 2011 [a]).

2.2.2.2 Fertilizing the Soil

The most important nutrients to crop enhancement found in fresh manure are nitrogen and phosphorus. Plants rely on essential nutrients such as nitrogen and phosphorus, in their inorganic form, for healthy growth (USEPA, 2002). Fresh manure contains both organic and inorganic forms of nitrogen with 60 to 90 percent of the total found in the organic form (USEPA, 2002). Organic nitrogen is normally found as urea whereas inorganic nitrogen is in the form of ammonium and nitrate. The nutrient can transition between organic and inorganic forms via the soil-atmosphere nitrogen cycle.

Commercial nitrogen is generally applied to fields in the inorganic form of nitrate or ammonium. Nitrogen in this form is highly water soluble and thus easily available and usable by plants (Discovery Farms, 2011). However, existing in the soil water means it can be removed easily from the field via runoff from heavy rains or snowmelt in the spring or by leaching into the groundwater. The potential for short residency of this nutrient form implies timing of the nutrient application should be just before the plants need it. However, around 95% of crop's nitrogen need occurs after the plants, such as corn, are too tall to allow spreading by heavy equipment (UM-Extension, 2012 [b]). Farmers applying commercial nitrogen take a risk that some of their nutrient investment will not be available in the soil by the time their crops need it the most.

Augmenting fields with manure as a source of nitrogen has several advantages. The obvious advantage is it provides a necessary method of manure disposal. It can also be a less expensive form of fertilizer thus reducing farming costs and helping with an operation's bottom line. As mentioned earlier, rising commercial fertilizer costs have transformed manure into a valuable commodity that can be used to supplement nitrogen and phosphorus in fields, but also can be sold to neighboring farms for income. Manure provides nitrogen in a mostly organic form. The organic form is more stable than the inorganic form and acts like a slow-release version of nitrogen fertilizer (UM-Extension, 2011 [a]). As long as soil temperature remains above 50 degrees Fahrenheit, soil microbe activity will act on the organic forms of nitrogen converting them to the plant-accessible inorganic forms (Frame, 2011). This process allows farmers to apply manure earlier in the spring, or even during the winter, ensuring the slow transition from organic to inorganic will result in better timing of nutrients for crops in later spring and early summer (UM-Extension, 2011 [a]).

The other major nutrient found in manure that is essential for plant growth is phosphorus. It is found in organic and inorganic forms in manure and, like nitrogen, the organic form dominates the percent of total phosphorus in manure (nearly 70%) (USDA, 1992). Similarly to nitrogen, phosphorus also breaks down over time to inorganic phosphate compounds which are usable by plants (USEPA, 2002). A key difference between the two nutrients is that phosphorus is not soluble in the soil water but is attached to soil particles (UM-Extension, 2012 [a]). Therefore leaching into the ground water or water draining from fields is not the major concern for phosphorus losses. Instead the greatest focus is on events where soil is eroded and carried off the fields due to overland flow. This mechanism carries off the soil particles to which the phosphorus is attached, thus transporting the nutrient into nearby water bodies (UM-Extension, 2012 [a]).

To maximize nutrient availability in their fields and cost-benefit of applying manure, farmers need to develop a keen interest in the timing of their manure spreading operations. As mentioned earlier, manure is not just a waste by-product anymore. In order to maximize their reduction in commercial fertilizer costs farmers must maximize their acreage covered by manure resources. However, that is only half of the process. That manure needs to remain on the fields long enough to breakdown into inorganic nitrogen and then stay in place to be available when the crops require it. It does not benefit the farmer if the manure and associated nutrients are transported from the fields before that can occur. Therefore farmers need to wisely choose fields which have the highest probability of holding that manure as well as select the right time to apply so that it does not runoff before it can break down and be incorporated into the soil. The first choice, selecting the locations, is a task that can be studied over time, changes very little over time, and can be documented in their nutrient management plan. The second choice, selecting the right time to apply, is a more dynamic and challenging decision to make. It is also the driving motivation for the RRAF tool which aspires to help farmers make the best decision they can regarding timing of manure application.

2.3 Environmental Impacts of Manure Contaminated Runoff

Although there are many advantages to spreading manure on agricultural lands, it is also important to recognize there are some serious risks associated with the practice. Specifically, the risk involves improper introduction of pollutants from the manure into the environment. There are several ways contamination can occur: direct discharge to surface water, spills during recovery or transport, and leaching from stacked piles or containment structures into ground water. Another major type of incident involves fresh surface-applied manure being carried off fields by runoff from rainfall or snowmelt. The RRAF focuses on the final incident example above by trying to minimize the occurrence of contaminated runoff events by warning producers of future high-risk runoff situations.

Incorporated into fields, nitrogen and phosphorus from manure are valuable assets. However, when those nutrients are displaced from the fields they become serious pollutants in nearby water bodies. The nutrients are not the only ingredients of consequence in manure. The EPA (USEPA, 2002) lists other components and pollutants of concern found in manure: organic matter, solids, pathogens, odorous compounds, salts and trace elements, and even antibiotics, pesticides, and hormones. The EPA goes on to clearly state that animal manure is the primary pollutant from animal feeding operations (USEPA, 2002) and overall agricultural operations are the leading contributor of impaired water quality in water bodies across the U. S. (USEPA, 1990). Of significant concern to the EPA and local regulatory agencies is the industry trend of increasing number of CAFOs which have an increased potential for environmental impact. These animal dense operations produce a lot of manure that usually ends up spread over a concentrated localized area, especially if there are many CAFOs in a regional area. If poor manure application decisions are made by these operators the scale of environmental impacts can increase in severity due to the increased magnitude of manure applied and available for transport to adjacent water bodies (USEPA, 2002).

Once manure moves from fields to water bodies the varying impacts observed will depend on the specific pollutant in question. Impairments in water quality range from environmental degradation to adverse health effects for humans. The scale of impacts is also variable depending on how often contamination occurs and how much pollutants and nutrients enter an aquatic system at a given time. Possible impacts could range from a temporary decrease in water quality due to a localized surface runoff event, to more severe degradation caused by consistent influx of nutrients from poorly managed manure applications in neighboring fields (USEPA, 2002).

The easily soluble nitrate form of nitrogen can be carried into streams from runoff or leached through saturated soils into ground water supplies. Once in drinking water supplies, increased nitrate levels can lead to higher water treatment costs and risks to human health (USEPA, 2002). The EPA identifies nitrate as the most widespread contaminant from agricultural sources in drinking water wells across the country (USEPA, 2002). An estimated 4.5 million people relying on wells are exposed to increased nitrate levels due to manure pollutant contamination (USEPA, 1990).

Although nutrients occur naturally in the environment and are essential for plant growth, when excessive amounts are introduced into an ecosystem, such as a small water body, eutrophication can occur. In aquatic ecosystems phosphorus is known as the limiting nutrient for aquatic plants. When excess phosphorus enters streams and lakes, there is a dramatic increase in plant and algae growth which can rob the system of other nutrients and decrease light infiltration (UM-Extension, 2012 [a]). Algal blooms can also reduce available oxygen in the water, clog treatment plant intakes, and cause undesirable tastes and odors to the water (USEPA, 2002). Compounding the increased phosphorus effect is the typical nitrogen to phosphorus ratio of 2 - 3 to 1 in manure. Field crops require a ratio closer to 4 – 9 to 1 (N/P). Since manure is applied based on nitrogen requirements, up to 3 times the needed amount of phosphorus is being added to soils in agricultural fields (USEPA, 2012

[b]). The extra amount of phosphorus physically cannot be used and therefore is highly likely to be transported away from the fields.

Eutrophic environments combined with excess nitrates encourage proliferation of toxic organisms and bacteria which can harm people via contact and kill fish (USEPA, 2002). In fact the most visible, and often dramatic, impacts of manure on aquatic ecosystems are large scale fish kills. Either by a sudden, massive influx of manure overwhelming the system or by slower development of degraded water quality in these water bodies, fish kills can be a serious issue (USEPA, 2002). States are not required to log fish kills, yet, the EPA was able to develop an incomplete dataset covering 19 states for most of the 1980s and 1990s. They tallied close to 6,000 fish kill events with an estimated 157 million fish killed. Focusing on Wisconsin for the years 1988 – 1998, there were 70 events recorded and an estimated 170 thousand fish killed. That averages out to around 6 events per year with around 2,400 fish killed per event (USEPA, 2002). Long duration degradation has also been found to decrease biodiversity in streams. Streams located downstream of large AFOs in Indiana were observed to have fewer fish and fewer fish species than control streams in the region (USEPA, 2002).

In their 2000 report, the EPA (USEPA, 2000) indicated that agricultural activities were the leading contributor of water quality impairments for rivers, streams, lakes, ponds and reservoirs and the fifth leading contributor for estuaries across the nation. These contributions lead to the impairment of 129,000 river miles, 3.2

million lake acres, and over 2,800 square miles of estuary across the U.S. Agricultural activities were found to affect water quality on rivers, lakes, and estuaries in 48, 40, and 14 states respectively (USEPA, 2000).

All of these excess nutrients make their way through the river systems and are funneled to the Gulf of Mexico or the Great Lakes. A significant environmental impact is found in the increasing size of the hypoxic zone of the northern gulf. Hypoxic zones occur naturally and are found all over the world. The northern Gulf of Mexico (NGOM) hypoxic zone peaks each summer and is found to be caused by nutrients delivered from the Mississippi River and natural stratification due to freshwater and seawater interactions. The NGOM hypoxic zone has been shown to be steadily increasing in size during its summer areal peak since first documented in 1972 (USEPA, 2012 [a]).

It has long been known that the hypoxic zone is reinvigorated each summer by nutrients loadings, specifically nitrogen and phosphorus, most commonly caused by agricultural operations in the watershed (Mississippi River Gulf of Mexico Watershed Nutrient Task Force, 2011). It has been shown that 38% of the total nitrogen load and 26% of the total phosphorus load is sourced from the Upper Mississippi River watersheds. These totals are second highest behind the Ohio River watershed for both nutrients (Mississippi River Gulf of Mexico Watershed Nutrient Task Force, 2008). Nutrient loads delivered to the gulf are dependent on the precipitation and corresponding runoff that is generated in the source watersheds. Studies have shown that nutrient loads measured in May are the most important as they correspond with the timing of the peak development of the hypoxic zone due to travel times to the gulf (Mississippi River Gulf of Mexico Watershed Nutrient Task Force, 2011).

The Hypoxia Task Force is renewing their goal of improving and expanding nutrient management by including two new statements in the Gulf Restoration Strategy (Mississippi River Gulf of Mexico Watershed Nutrient Task Force, 2011). The statements in the 2011 annual report are as follows:

"(1) hypoxia, fed by nitrogen and phosphorus pollution in the MARB (Mississippi/Atchafalaya River Basin), is an obstacle to a healthy and sustainable Gulf ecosystem, and (2) the key to curbing oxygen starvation in the Gulf is effective nutrient management planning and targeted implementation of best management practices in upriver and coastal priority watersheds."

In addition, for 2012, the Task Force planned on supporting state nutrient strategy development, enabling collaboration and ensuring "lessons learned" are shared among member states and federal agencies, and finally identifying opportunities for obtaining financial and technical support (Mississippi River Gulf of Mexico Watershed Nutrient Task Force, 2011). It appears that the RRAF product meshes well with both Wisconsin water quality goals as well as with the national level task force goals and mission.

2.4 Review of Manure Application Guidelines & Regulations

The risks associated with applying manure to crop lands suggest the need for proper guidelines, best management practices (BMPs), and regulations to reduce opportunities for nutrients and pollutants to runoff into adjacent water bodies. Local, state, and federal agencies provide guidance, suggestions, and rules for producers to follow while spreading manure on their fields. This section will review some of these regulations specific to Wisconsin.

The RRAF is focused on reducing contaminated runoff leaving farm fields by helping producers avoid applying during risky soil and weather conditions. Therefore, this discussion will center only on rules and guidance that relate to the aspect of manure application. More specifically, the emphasis will be on guidance relating to environmental conditions during application as well as the timing.

A vast majority of the written documentation regulating manure and nutrient management is geared towards planning and more static aspects of the process instead of the more dynamic timing decisions. Guidance for evaluating, tracking, and planning nutrient requirements for fields based on crop rotations is one example. Another topic includes rules describing the “where” component to application such as selecting fields with minimal slope, ensuring conservation practices are in place, and identifying drinking wells, tile inlets, and surface water pathways. These two examples highlight tasks that can be done well ahead of the application and do not change over the farm year. Manure management issues based around manure storage overflow, stacked manure piles, or runoff from farm operation areas (feedlots, holding areas, etc.) also do not represent the primary audience of the RRAF and therefore will not be discussed in this review.

2.4.1 Regulations for Manure Application

Like many other states, Wisconsin has instituted a variety of rules and regulations for the agricultural sector to follow regarding nutrient management and conservation practices. Under these state codes, every farmer must comply with NR-151 which sets performance standards and details prohibited actions for farms (WDNR, 2004) (Wisconsin Administrative Code, 2012). Wisconsin administrative code ATCP-50 further identifies conservation practices that farmers must follow to meet performance standards set in NR-151 as well as defining requirements for Nutrition Management Plans (NMP) (Wisconsin Administrative Code, 2011). All farms that apply manure or commercial fertilizer must implement a NMP for all croplands (DATCP, 2012 [a]). Further, these NMPs must be written to meet the Natural Resources Conservation Service (NRCS) technical standard 590 (WDNR, 2004). Additionally, larger operations such as CAFOs, and some smaller AFOs designated by the state, are required to obtain a Wisconsin Pollutant Discharge Elimination System (WPDES) permit which requires a 590-approved NMP and to follow land application requirements set in NR-243 (WDNR, 2012 [b]).

With the exception of WPDES operations, there is a major caveat with regard to the enforcement of the adoption of NMPs. Reality is that farms cannot be required to create and follow a NMP unless at least one of the following conditions was met: (1) the farmer accepted cost-share grants to implement nutrient management or install manure storage, (2) the farmer is participating in the Farmland Preservation Tax Credit program, (3) local manure storage ordinances require it, or (4) the operation was found to cause contaminated discharge to water bodies (DATCP, 2012 [b]). Further, there have never been enough state funds to offer every farmer cost sharing that qualifies (Jenks, 2012). Therefore regulators are forced to focus their efforts to provide cost sharing funds to the most critical farms. The MMTF proposed the assumption that if 25% of the state's farmland was designated as needing protection, $7 - $14 million would be needed annually for 5 – 10 years to provide cost share funds to implement NMPs on these threatened or degraded lands (WDNR & DATCP, 2006). Although farmers can voluntarily create and follow a NMP, reality is that there are many farms across the state that still do not use them.

The state of Wisconsin strives to make the process of creating and implementing a NMP as easy as it can via free software and web tools. In 2012 the state received over 3,800 NMPs which covered over 1.9 million acres. This represents about 22% of Wisconsin cropland covered under NMPs (DATCP, 2012 [b]). DATCP has seen over a 20% increase in the number of acres under NMPs since 2010 as well as a steady increase annually in the number of acres since 2006 (DATCP, 2011).

The preceding discussion has revolved around state codes and NMPs. Some clarification on what is meant by NMPs could be beneficial. NRCS code 590 (NRCS,

2005) defines nutrient management as managing the amount, source, placement, form, and timing of the application of nutrients and soil amendments. It goes on to describe the purpose is to establish acceptable criteria and documentation requirements for applying and budgeting of nutrients for plant production. The stated goal is to minimize nutrient entry into surface water, ground water, and the atmosphere while maintaining and improving the physical, chemical, and biological condition of the soil (NRCS, 2005). The state of Wisconsin expands the NMP definition by including additional wording stating a NMP is a crop practice record reviewed annually and updated as necessary. The NMP must be prepared by a qualified planner and account for all nitrogen, phosphorus, and potassium nutrients applied to each field (DATCP, 2012 [b]).

Wisconsin DATCP also states that farmers implementing 590 standard NMPs are using one of the best management practices for reducing pollutants from water supplies. In addition, following these plans help farmers allocate their nutrients in the best way economically by not over applying and wasting their manure resources (DATCP, 2012 [b]). Having a NMP in place when a manure runoff mishap occurs doesn't protect against liability, but it will be better than if one was not being followed (DATCP, 2012 [a]).

The state regulations NR-151, ATCP-50, NRCS standard 590, and NR-243 will be reviewed below to highlight only the aspects that apply to the timing of manure application, the focus of the RRAF product. To obtain a full understanding of the complete rules and guidelines defining where and how much to apply is needed, the full regulations should be reviewed separately.

Wisconsin's version of NRCS code 590 (NRCS, 2005) has a few general references that apply to timing of manure application operations. The first reference found in section V-A-1-k, "Nutrient Criteria for All Sites", simply states that manure and fertilizers shall not run off a field during or immediately after application. In the second reference, "Nutrient Application Prohibitions" (V-A-2-a-4), the standard describes spatial restrictions but also introduces a rule where manure must be incorporated within 72 hours. The 72-hour window for incorporation (mixing manure with topsoil) is a common guideline and is used by DATCP in the RRAF product definition that will be discussed later in this paper. In section V-A-2-b guidance is defined on what to do if soils are frozen or snow-covered. In this situation, further spatial and landscape restrictions are introduced as well as a limit of 7,000 pounds of liquid manure per acre. The final required reference to timing or soil conditions by code 590 is found in the third section, "Nutrient Application Restrictions" (V-A-3-a). This guideline states no applications of unincorporated LDM are allowed on non-frozen, saturated soils. Section VI, "Considerations", part G states that the producer should consider delaying surface application of manure if precipitation capable of producing runoff is forecast within 24 hours of the planned application. This suggestion points right towards the goal of the RRAF and leads to why the RRAF can be successful. The RRAF takes into account the precipitation, soil conditions, and runoff potential and helps the producer make the best decision on when to apply.

Chapter Natural Resources NR-151 of the Wisconsin State Administrative Code references manure application indirectly in section 151.07 part (3) where it states all manure must be applied in conformance with a NMP which should be designed to limit nutrients from entering the waters of the state (Wisconsin Administrative Code, 2012). Much like NR-151, state code chapter Agriculture, Trade, and Consumer Protection (ATCP) 50 details many aspects of manure management for the state's producers (Wisconsin Administrative Code, 2011). However, there is little which is directly associated with timing of manure applications. One reference found was in subchapter VIII, section 50.62 (5), parts c and d. Here the code states that no cost sharing funds can be used to reimburse costs for manure storage systems unless manure is incorporated into the soil within three days after application and no manure from that system is applied to frozen or saturated ground.

The final state code reviewed will be chapter Natural Resources NR-243 which only applies to CAFOs and other medium sized operations that are regulated under WPDES permits (Wisconsin Administrative Code, 2007). Generally NR-243 is similar to the other regulations with the exception that it goes into much higher detail in what is allowed during winter applications. Section 243.13-2-b includes some of the requirements that must be followed during any land applications of manure. Timing applicable rules include that no manure may run off the field during application or be applied to saturated soils. An important difference from earlier regulations is found in part 4 of this section of NR-243. It states no manure may run off the field due to precipitation or snowmelt except in situations where the producer has complied with all land application restrictions and the rain event is equal to or greater than a 25-year, 24-hour rain event. More in-depth rules continue with part 11 stating manure may not be applied when snow is actively melting and water is running off the field. Time tolerances are also tightened in part 12 which states where required by standard 590, incorporation of manure must be accomplished within 48 hours of application. And finally, part 13 explicitly states no manure may be applied when precipitation capable of producing runoff is forecast within 24 hours of the time of application.

The more specific winter restrictions of manure application on frozen or snow covered ground begin in section 243.13-5 for solid manure and 243.13-6 for liquid manure. The regulation defines frozen ground to mean anytime the soil is frozen between a half inch and eight inches from the surface. For solid manure, surface application is allowed on frozen soil if the site is not prohibited by other reasons and the producers follow guidelines provided for slope and tillage practices. The same rules apply for snow covered ground. Beginning in 2008, a high-risk runoff period was introduced with more stringent restrictions. The high risk period lasts from February 1st through March 31st and applies if either of the following conditions exists where the manure is to be applied: snow depth of 1 inch or more, or the ground is frozen.

Surface application of liquid manure on frozen grounds is prohibited. Manure may be injected or immediately incorporated on frozen ground as long as there is not more than 4 inches of snow present. On grounds with less than one inch of snow cover, liquid manure application is allowed. If the ground is covered by one to four inches of snow, surface application is prohibited and only injection or immediate incorporation is allowed. On areas where more than four inches of snow exists only injection is allowed. During the high-risk runoff period defined above, the application of liquid manure is prohibited.

Reading through section 2.4 hopefully shows that understanding the impact of weather on the dynamic decision making process of when to apply manure is a key component to successful manure management. This is especially true in the winter season. Some regulations are in general context, while others such as NR-243 are more specific.

It is assumed that most farmers are aware of current soil conditions as well as what the forecast weather is to some degree. To reinforce the concepts and BMPs in the regulations farmers are reminded of the weather impact through countless web articles, research, and outreach programs across the country as well. Some web article examples include a University of Minnesota Extension article that includes a small reference about recording past and future weather conditions in a log to help with application record keeping and to aid the producer in the case of a runoff event (UM-Extension, 2011 [b]). An Iowa State article concludes with the reminder to be on the lookout for predicted rainfall or warming conditions that could cause runoff (Brenneman, n.d.) . Another example to highlight the importance of weather is provided by Fleming and Fraser (2000) who completed a review of 11 studies conducted from the 1970s to 1990s on winter manure spreading across the northern United States and southern Canada. In terms of nutrient loss, they concluded that the single greatest risk factor associated with manure application in the winter was the weather.

The discussion above introduces and solidifies the concept that paying attention to forecast weather conditions is important. However, the documentation reviewed provided very little information on how to transition from merely paying attention to the weather to making decisions based on it. The only reference to any real guidance came from the EPA in an appendix to their CAFO manual (USEPA, 2012 [b]). This reference was essentially instructions on how to obtain ensemble forecast precipitation guidance from the National Weather Service. Their recommendation is to focus on the 24 hour probability of precipitation over their area. If the probability is greater than 70% for either 0.5” or 0.25” value (depending on the producer’s judgment of soil conditions), then the producer should not apply.

Although Chapter 3 will include a more comprehensive discussion on how the RRAF is created, highlighting some of the components of the product now will showcase how it incorporates and addresses many of the regulations and BMPs that were just reviewed above. To start, the RRAF provides warning for future runoff derived from precipitation and/or snowmelt 10 days into the future. By tracking the soil moisture, the RRAF accounts for times when runoff could be generated from 0.10" of rain or indicate if 1.0" of rain is needed. Modeling the snowpack and including future temperatures helps with the snowmelt aspect that producers need to watch for. In addition, saturated soils are accounted for as RRAF runoff risk requires that condition by definition. A 72-hour future window is applied to a given day's risk value which is a way to include the 72-hour or fewer requirements for incorporating manure. A scenario highlighting the usefulness of the 72-hour risk window could include a time when a producer sees great conditions to spread on day 1 and plans to incorporate on day 3. However, the producer is not aware of high runoff potential on day 3 which results in contaminated runoff leaving his fields. The RRAF is built to avoid these scenarios.

By incorporating the entire soil-atmosphere system and providing forecast conditions out ten days it is hopefully clear to see how the first-of-its-kind RRAF product is useful to producers, regulators and environmental monitors. Adoption of the product will help producers manage the timing of applications to ensure they get the maximum nutrient value of their manure by keeping it on their fields and out of the state's water bodies.

2.5 Seasonal Manure Applications and Field Scale Runoff Characteristics

This section will review some of the advantages and disadvantages associated with applying manure in different seasons of the year irrespective of regulations. In addition, characteristics of edge-of-field (EOF) runoff in relation to nutrient management will also be discussed. Much of the field scale runoff research reviewed comes from the University of Wisconsin-Madison Discovery Farms program. This program conducts field research on real working farms across Wisconsin to identify effective environmental management practices that can coincide with a profitable agriculture industry (Discovery Farms, 2012 [a]). In addition, four Discovery Farm edge-of-field observed runoff datasets were used in the validation of the RRAF model. Detailed descriptions of the four observed datasets are available in section 3.5.

2.5.1 Spring Application

Application of manure is typically thought to best be applied in late spring and early summer as close to the time of planting as possible. This timing helps to maximize crop nutrient uptake and minimize nutrient losses due to runoff or leaching (UM-Extension, 2012 [c]). However, it is commonly known that spring is a very risky time to spread as fields are either covered in snow or remain fairly saturated from the snowmelt. Discovery Farms research (UW-Extension, 2011 [a]) confirms this and extends the risky period into June due to the combination of the soil conditions and an increase in rainfall across the Midwest at this time of year. Factors associated with crop growth also inflate the risk during the spring and early summer. Newly planted seed provides minimal canopy cover by this time to shield soil from raindrop impact and has minimal water demands to help pull moisture from the soil column (UW-Extension, 2011 [a]).

The advantages of spring application generally outweigh the risks just listed above. The decreased duration of nutrient exposure on the fields leads to less potential for losses and more nutrient benefits for the crop. A study in Minnesota illustrated the latter point when the authors found that manure injected in April resulted in 5% higher yields than manure injected in the fall (UM-Extension, 2012 [c]). However, hectic farming operations in the spring often force producers to schedule manure applications at other, less ideal, times of the year. Due to this reality, fall and winter spreading has become more common in the Upper Midwest and introduce their own set of advantages and risks.

2.5.2 Fall Application

An obvious advantage of fall applications is from the lack of competition with other required farming operations. Planting and harvesting activities have been completed and more time and labor are available for handling manure distribution. Over the summer months, fields full of crops lead to producers having to fill storage facilities with manure until harvest occurs. Therefore, they can be anxious to start applying in the fall to decrease their stored manure before the winter. This helps them avoid or minimize the need to build much larger manure storage structures that could handle the entire fall and winter manure production (USEPA, 2012 [b]). Although it stretches the definition of fall, Discovery Farms (Frame, 2011) have shown that August and September have very low potential for runoff from their field studies. This time of the year could be a good time to spread, strictly in terms of losses from runoff, if fields are available.

2.5.3 Winter Application

As alluded to in the regulation summary, winter applications introduce the most challenges and risk for losing nutrients from croplands. There are, however, some advantages to this time of year. Much like fall applications, spreading in the winter is usually easier on a producer's schedule. In addition, frozen soils are more capable to withstand heavy loads from spreading equipment without causing compaction of the soils.

However frozen and snow covered ground enhances risk by its very nature. With frozen ground, the soil surface impacts infiltration to make it either impossible or incredibly dampened. The second condition introduces more potential runoff volume, via the snowpack, which could increase the potential transport of nutrients from fields. Another disadvantage of winter application is that incorporation is not always possible, or allowed, due to frozen ground or a deep snow pack. Not incorporating results in the manure remaining on the surface allowing the small percentage of inorganic nitrogen in manure to volatize away quickly and be lost. The manure that remains is obviously sitting on the top of snow and again more susceptible to transport off the fields (UM-Extension, 2012 [c]). Discovery Farms has found that the majority of annual nitrogen losses are due to applications on frozen soils (Discovery Farms, n.d.). The losses are in the organic form mostly due to the cold temperatures limiting the conversion to nitrates.

There have been many studies that investigate the impacts of winter spreading and provide some general suggestions to accomplish this safely. The following are some examples of BMPs for winter spreading: (1) try to spread early in the season before a significant snowpack has accumulated, (2) avoid spreading over a deep snowpack late in winter, and (3) if applying is needed in late winter, let most of the snowpack melt off first (Brenneman, n.d.). Another suggestion is to choose fields that just came out of soybean production instead of corn. The taller corn stubble will trap and hold a deeper snowpack. This leads to more runoff in the spring and has been shown to increase nutrient losses (Lorimor, 1995).

2.5.4 Edge-of-Field Runoff Characteristics

As mentioned, Discovery Farms, in partnership with the United States Geological Survey (USGS), has set up many monitoring stations on farms in Wisconsin. This equipment is usually set up to monitor runoff, sediment, and nutrient loads from agriculture fields to evaluate the effects of the representative landscape, soils, and farming practices typical in Wisconsin on minimizing nutrient losses and improving crop productivity (Stuntebeck et al. 2011). Over the years they have shared their findings and insights into nutrient management and runoff via the Discovery Farms website and newsletters to their sizable audience in the state. Many of the Discovery Farms references cited here are derived from findings included in the Stuntebeck et al. (2011) report. USGS data collection methods at Discovery Farms are provided in greater detail in a USGS report written by Stuntebeck et al. (2008).

The Stuntebeck et al. (2011) report summarizes hydrologic and water quality observations on 23 edge-of-field stations across 6 farm sites for 2003 to 2008. Of these 23 stations, four were used in the development of the RRAF. During their study period mean annual precipitation was 32.8 inches which is about 3 percent lower than the 30-year mean. The mean annual runoff for all of their stations was 2.55 inches, which is about 8 percent of the precipitation. There was variability in the runoff efficiency over the study period where less than 1 percent of the annual precipitation was measured at some farms in 2006, while in 2008 a couple farms reported nearly a 15 percent rate. The study also pointed out that annual runoff was not simply a factor of annual precipitation. They highlighted rainfall intensity, soil moisture, and existence of frozen soils as key contributors to runoff generated.

In terms of runoff, two important time periods were identified in this study. The first was during frozen ground conditions (February and March), and the second window occurred during non-frozen ground conditions in May and June. February and March observed the highest mean monthly runoff (0.41 and 0.87 inches) which accounted for 50 percent of mean annual runoff. The percentage of precipitation that was observed as runoff was 28 and 39 for February and March, respectively.

Although only 11 percent of the annual precipitation fell during this time, the possibility of storing precipitation of previous months in the snowpack could have occurred and impacted those percentages. Runoff was caused mostly by snowmelt or rain on snow events during this time period and was longer in duration, greater in volume, and lower in peak discharge than events occurring during non-frozen ground periods. Also noteworthy is that only March observed runoff at every monitoring station during every year studied. January, February, and April were next highest in terms of runoff frequency with a 50 percent score whereas August, September, October, and November had the lowest frequencies of near 20 percent of the time. Overall, the portion of mean annual runoff that occurred during frozen ground periods was 54 percent while the portion during non-frozen ground periods was 46 percent (Stuntebeck et al. 2011). Substantial variability was observed, however, as there were sites for a given year where 100% of their annual runoff occurred during frozen ground or snow covered conditions (Discovery Farms, n.d.), while other sites recorded 70% of their annual runoff during non-frozen ground conditions in other years (Stuntebeck et al. 2011). Overall, their study did find that for 9 of their 26 datasets, nearly 100 percent of annual runoff was during frozen ground conditions and at least 50 percent of annual runoff was recorded during these winter conditions in 16 of the 26 datasets.

The second important period was during non-frozen ground conditions in May and June where these two months accounted for 31 percent of the annual runoff on average and recorded mean monthly runoff of 0.32 and 0.48 inches respectively. Runoff during these periods was typically due to high-intensity rainfall or smaller rainfall events combined with high soil moisture conditions. Minimal crop cover during this time was also a factor. Overall, the months February through June are the most important months in terms of runoff where the months of February, March, May, and June contributed 81% of the average annual runoff. No other individual month recorded over 5% of the total (Discovery Farms, n.d.). Ninety percent of mean annual runoff occurred over the time span of January through June (Stuntebeck et al. 2011).

With respect to nutrient concentrations and yields, Stuntebeck et al. (2011) noted the following observations. During non-frozen ground conditions, sediment concentrations and yields were significantly higher than frozen ground conditions. Ninety percent of suspended sediment yield was during non-frozen conditions and over 80 percent was in May and June alone. Phosphorus yields were also higher in non-frozen ground conditions with 60 percent observed and the highest yield months corresponding with the highest runoff months (February, March, May, and June). Total nitrogen yield was generally evenly split over frozen (52 percent) and non-frozen ground conditions (48 percent) with the highest yield months aligned similar to those of the highest phosphorus yields and runoff. In addition, mean monthly yields of total nitrogen and total phosphorus were strongly correlated suggesting the source of both were likely the same.

In terms of mean event concentrations, suspended sediment and phosphorus were higher in non-frozen conditions while total nitrogen concentrations were higher during frozen soil conditions. For both nitrogen and phosphorus, the highest event concentrations occurred when manure was applied right before runoff (days to weeks) during frozen ground conditions and when sediment concentrations were highest during non-frozen ground conditions (Stuntebeck et al. 2011).

Many years of accumulating runoff and water quality data across the state has allowed the Discovery Farms team to conclude that at the field level there were consistently one or two major runoff events in every year that dominated in importance with regards to nutrient runoff. These events usually accounted for nearly half the annual nutrient losses and generally contributed a large percentage of the annual runoff volume (UW-Extension, 2011 [b]). Interestingly, these major events occurred at various times of the year. For some sites, they occurred in May and June others were due to snowmelt on frozen soil, while another was in fall during a heavy rainfall event (UW-Extension, 2011 [b]). The promoted take home message outlining the vulnerable conditions susceptible to weather impacting nutrient management includes two common themes found in this section: (1) sparse ground cover during spring or fall, and (2) snow melting on frozen soil. A third situation not mentioned much to this point involves times when rainfall exceeds the design criteria of the management practices in use (Discovery Farms, 2011).

This data has also shown that nutrient losses from fields positively correlate with the volume of runoff from that same field. It was also discovered that the shorter the amount of time between manure application and runoff occurring increased the potential for nutrients to be lost (Discovery Farms, n.d.). Sediment loss was almost exclusively tied to non-frozen ground conditions which, not surprisingly, also corresponded to the time when most of the phosphorus was lost from the test fields as most phosphorus is bound to soil particles. Conversely, most nitrogen losses were during the frozen ground periods (Frame, 2010).

Focusing further on manure applications during frozen ground conditions, Komiskey et al. (2011) detailed nutrient and sediment in runoff from three adjacent fields during four consecutive winters where LDM and solid-beef manure (SBM) were both applied in typical fashion by the farm operator. The fields were all 40 acres or less, no-till managed, and located on a Discovery Farms site in southwestern Wisconsin. The time period studied consisted of the winters of 2003 – 2004 to 2006 – 2007 where frozen ground periods typically occurred from mid-December to mid-March and runoff events were derived from snowmelt, rainfall, or a combination of both. Komiskey et al. (2011) reported that although each field received similar amounts of precipitation during the four frozen ground periods, the number and volume of runoff events varied greatly over the years. Further, runoff volumes were related closer to the timing, type, and intensity of precipitation as well as air temperatures and snow-pack conditions instead of strictly the presence of frozen ground.

Regarding nutrients carried in the runoff, in general, nitrogen and phosphorus concentrations and losses were lower if manure was applied in fall or early winter and the first runoff event did not occur for a couple months after application. On the opposite side of the spectrum, concentrations and losses were dramatically higher if manure was applied less than one week before the next runoff event, even if application rates were low. Their conclusion was that the timing and amount of LDM applied to fields during frozen ground periods were key contributors to the nutrient concentrations and losses. Therefore, manure management decisions during these conditions are especially important with regards to minimizing annual contaminated runoff (Komiskey et al. 2011). In addition, all factors being the same, it was also found that LDM application contributed more nutrients to runoff than application of SBM during the study.

Discovery Farms sums up their field scale runoff research by informing the producers of Wisconsin to focus on three major considerations besides accepted conservation practices. The first is to be especially aware of critical runoff periods, which they define as the snowmelt period (especially February and March). The second is to focus on when field conditions indicate soil moisture is reaching saturation level (over 30%). The third is to consider the timing of manure applications (Frame, 2011; Frame, 2010). They go on to stress that day-to-day decisions are just as important as long term conservation practices by the producer, and by following this guidance derived from real farmland research, reductions in nutrient contaminated runoff from their fields will result (Frame, 2010).

2.6 Exploration of Similar Runoff Risk Guidance Available

Impacts associated with contaminated runoff from agriculture lands ranging from degradation of the smallest streams and lakes, to millions of contaminated drinking wells, to the growing hypoxic zone in the Gulf of Mexico, reinforce the idea that a tool is needed to help producers manage the timing of their manure applications. Field scale research, such as the studies by Discovery Farms, has consistently supported the concept that to reduce nutrient losses, emphasis has to be placed not only on long term conservation practices, but also on the day-to-day dynamic decisions of when to apply. These decisions carry consequences more often than not, mainly due to what the future weather conditions hold in store.

Up to this point, producers were left with little guidance or instruction on how to gather and evaluate weather forecasts and process them to help make these critical timing decisions. Suggested procedures were often dealt entirely around forecast precipitation forecasts. Some guidance also mentions monitoring observed rainfall, trending towards an antecedent moisture condition type index as a crude way to ingest the soil moisture variable.

Strictly using precipitation-based forecast guidance is better than no guidance. However, previous discussion in this section highlighted many risk factors left out with this approach such as soil saturation, frozen ground, snowpack, and snowmelt only runoff. The RRAF decision support tool incorporates all of these risk factors and marks the first time producers across an entire state have the ability to use a real-time model that incorporates soil and meteorological forecasts to assess general risk for their farms.

A brief internet search, as well as coordination with DATCP partners, was accomplished to document other similar forecast guidance available for manure and nutrient application. A brief summary of the products found, valid in December 2012, is listed below.

The first similar product found is called the *Manure Spreading Index (MSI)* and was developed by the Oregon Department of Agriculture for use by CAFOs. This tool was developed in 2004 with the aid of a grant from the EPA. The grant is said to have expired but the site remains operational. This tool is an example of an antecedent moisture index combined with forecast precipitation sourced from the National Weather Service. They developed a formula based on the rainfall inputs to come up with a three category risk scale. This product can be viewed at this website: http://msi.jimlittle.net/index.shtml

The second example of a forecast tool was called the *National Weather Service Forecast & Farmers Map*. Little documentation was found but it seemed to be a joint product of a NWS Weather Forecast Office in Birmingham, Alabama and the Alabama Cooperative Extension System which is affiliated with Alabama A&M and Auburn Universities. Again, this product seems to be derived only from forecast precipitation. Instructions state that if the chance of rain is 50% or higher in the next three days, producers should consult the map product and are cleared to apply if their area is colored white. No further documentation is provided to define what the threshold is to place an area in either of the two risk categories. This product can be found at: http://www.srh.noaa.gov/bmx/adem/farmers_map.php

The Whatcom Conservation District, located in western Washington State, has introduced its *Manure Spreading Advisory* webpage. Documentation is not complete on the webpage, but some investigation suggests that this product is derived from only forecast precipitation. There is an accompanying spreadsheet tool that asks the producers specific questions to help provide a field specific risk assessment. However, their spatial map and four-category risk scheme offers no details on what determines the risk. Clicking on individual watersheds provides a popup box that lists one day and 3 day forecast rainfall totals. The look and feel of this tool is very similar to the RRAF. This group contacted the RRAF Advisory Group at the end of 2011 and was given a copy of the webpage code to help them develop its own tool. The MAS can be found at: http://whatcomcd.org/manure-spreading-advisory

Finally, the last of the similar guidance tools was a reference to a project under development by the U.S. Department of Agriculture – Agriculture Research Service

(USDA-ARS). Its goals are to create a webpage that issues 1-day and 5-day runoff forecasts for areas around the Chesapeake Bay area. This team is also using forecast precipitation and soil moisture data to look for correlations with runoff data to help them develop a model to predict runoff risk. The documentation found is over two years old and the current status of the project is not known. The website for this project can be found at: http://www.ars.usda.gov/is/pr/2010/100818.htm

The review of these similar products suggests an encouraging development in the realm of nutrient management. It appears technology is finally heading towards producing usable, easily accessible decision support tools that both producers and regulators have been requesting. At this time, it appears the RRAF is the most comprehensive tool with the largest spatial coverage and shortest update cycle. More details on how the RRAF was created will be discussed in Chapter 3. It is expected that the next several years will see both an increase in the number of tools across the country as well as an increase in the quality and accuracy of the guidance.

3. DATA AND METHODS

3.1 Spatial Scope of the RRAF Decision Support Tool

The entire state of Wisconsin is covered by the RRAF product by using 216 of the 1,173 sub-watersheds simulated by the NCRFC. This accounts for all basins located inside Wisconsin or being split by the state border. The basins are named via a NWS methodology using five character IDs called the National Weather Service Location Identifier (NWSLI) where the last two characters are state specific. For example, all basins centered in Wisconsin end with W3 since Wisconsin is the third U.S. state that starts with "W" in alphabetical order. The mean modeled watershed size for the RRAF is around 300 mi^2 which is slightly larger than the mean size of all the NCRFC basins (Table 1). The total area covered by the RRAF is 65,000 mi^2 which is about 19 percent of the entire NCRFC area of responsibility.

Table 1. Sizes of NCRFC lumped model sub-watersheds.

	mi^2	**km^2**	**acres**
RRAF Basins (216)			
Smallest	9	23	5,760
Largest	1,837	4,758	1,175,680
Mean	301	780	192,640
Std Deviation	283	733	181,120
Sum	65,003	168,357	41,601,920
All NWS Basins (1,173)			
Smallest	7	18	4,480
Largest	3,061	7,928	1,959,040
Mean	291	753	186,240
Std Deviation	281	728	179,840
Sum	341,357	884,110	218,468,480

3.2 Forcing Data for NCRFC Models

Forcing data required to drive the NCRFC models is broken down into two categories: temperature and precipitation in both observed and forecast timeframes. The standard daily model runs generally start 11 days in the past and run 14 days into the future. All forcing data, observed and forecast, is processed to compute 6-hourly mean areal averages for each sub-watershed.

Observed temperature data is derived from hourly reports from the network of automated weather stations across the region known as either Automated Surface Observing System (ASOS) or Automated Weather Observing System (AWOS). Generally, 3-hour reports are averaged to provide a 6-hour mean. The ASOS and AWOS network is also used for observed precipitation however this data is complimented with an extensive network of human observers in the Cooperative Observer Program (COOP). COOP reports include many meteorological conditions including 24-hour total precipitation. All daily precipitation totals are defined on a "hydrologic day", which runs from 12Z to 12Z which corresponds with the morning forecast model schedule which uses its initial time (T-zero) as 12Z.

The observed data is then used to generate 6-hourly mean areal basin values, referred to as Mean Areal Precipitation (MAP) (Figure 5) and Mean Areal Temperature (MAT). To help time distribute the COOP 24-hour total precipitation amounts nearby ASOS/AWOS or radar estimates can be referenced to accurately depict the timing over the day on the four cardinal time steps (00, 06, 12, and 18Z) used by the model.

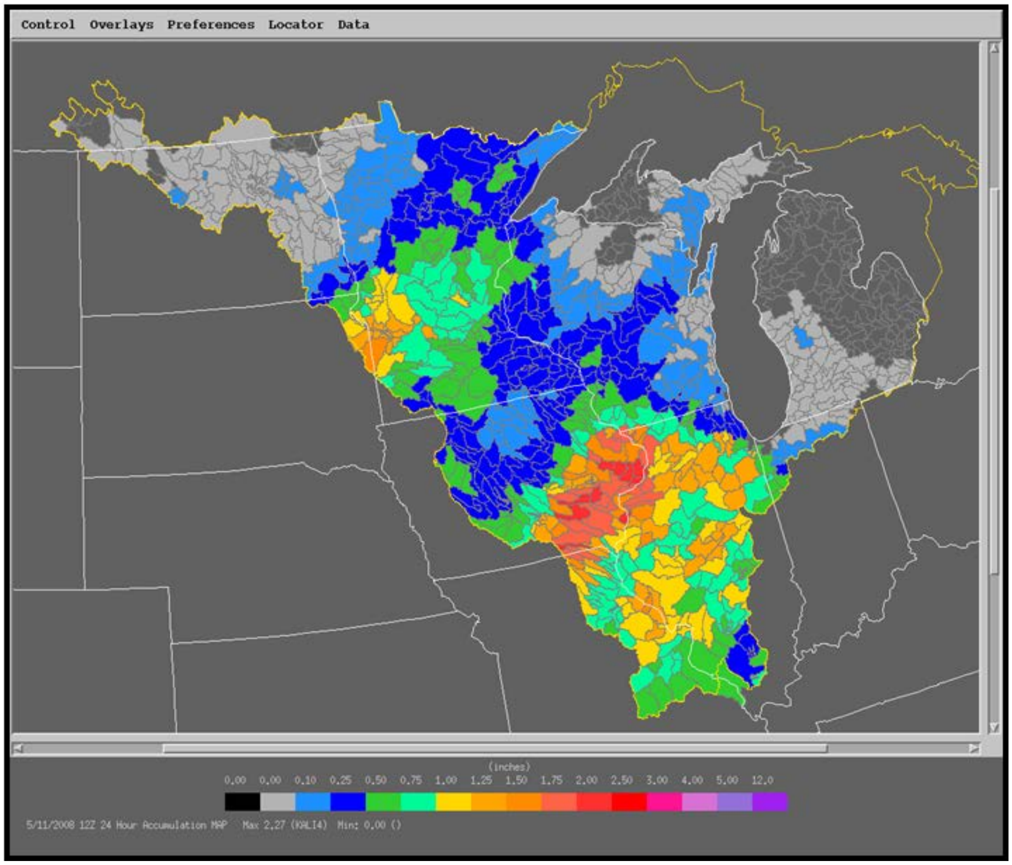


Figure 5. Example of 24-hour sum of 6-hourly mean areal precipitation (MAP) used as input into the hydrologic models. Source of data could be from either a gauge network or radars.

Quality control is a very important task which is assigned to one person daily. The gauge network and radar estimated precipitation need to be checked routinely for incorrect data which could influence the MAP or MAT values. Between the times of 10 – 14Z gauge data is continuously streaming in as observers report their daily values. Because of this the MAPs and MATs are routinely re-generated and ingested by the model every 10 minutes to ensure the most complete forcing time series is available. For the radar estimates, quality control is completed on the hourly radar-derived precipitation fields. These grids are automatically adjusted incorporating a bias derived from ASOS/AWOS hourly observations yet still need to be corrected. There are 27 NWS operated WSR-88D radars as well as 8 Canadian radars that cover the NCRFC watershed. A majority of the U.S. owned radars have completed their dual-polarization upgrade which is expected to provide better estimates of precipitation in the future (NWS, 2012). Examples of the observed gauge precipitation network and radar coverage are shown in Figures 6 and 7.

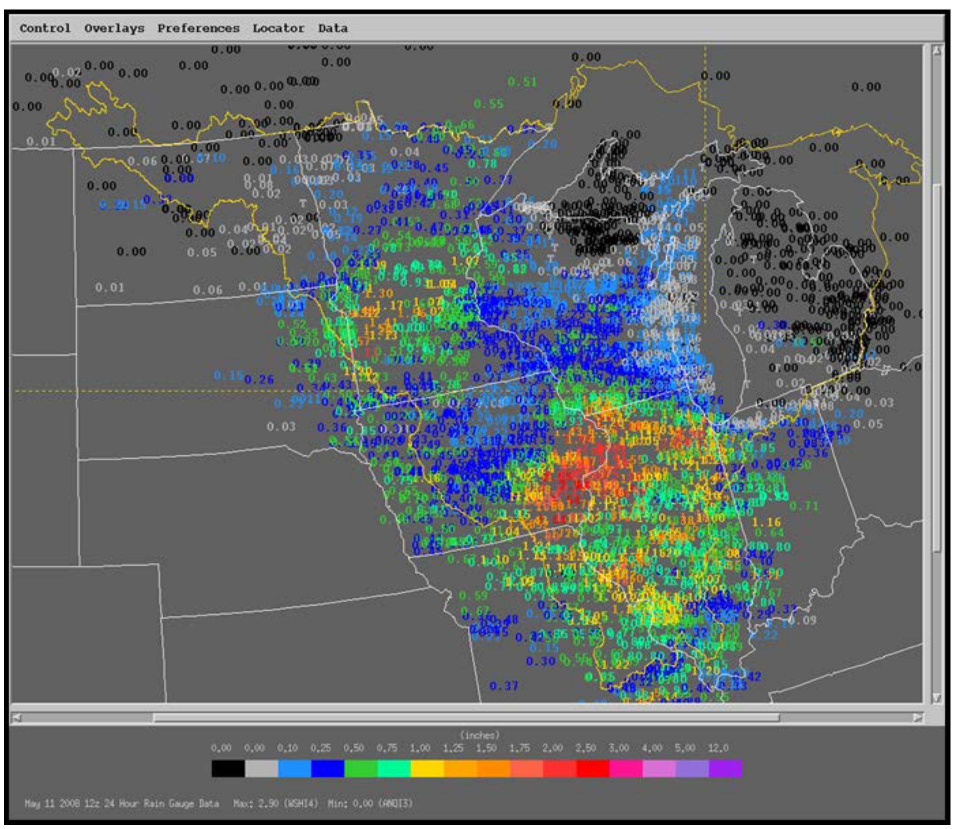


Figure 6. Example of observed precipitation gauge network across NCRFC region.

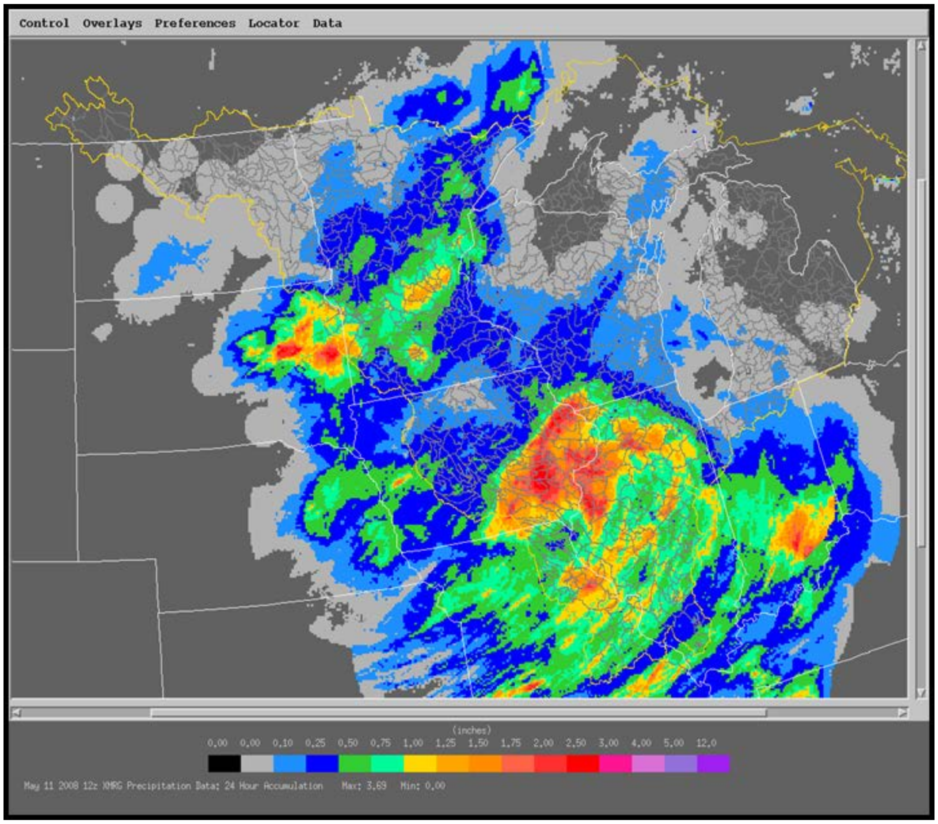


Figure 7. Example of radar derived 24-hour precipitation across the NCRFC region.

During operations the forecaster has the option to choose between the gauge network (the default choice) or radar derived MAPs on a basin by basin basis or for an entire Forecast Group. This is a common modification made by forecasters to compensate for poor precipitation estimates in one field or to match basin river gauge response. There are often times where using the radar field is preferred such as during the night when very few manual observations are available, or in sparsely populated areas where there is not a reliable gauge network to capture a weather event in its entirety. Conversely, the radar derived product has issues such as poor coverage in certain areas or bright banding due to melting snow or hail which increase reflectivity and inflate precipitation estimates.

The forecast component of forcing data for the hydrologic models is derived from several forecast model grids. Starting with forecast temperatures (FMAT), the first seven days are taken from the NWS National Digital Forecast Database (NDFD) with days 8 – 14 selected from the North American Ensemble Forecast System (NAEFS) ensemble mean. Forecast precipitation (FMAP) is generated using 5-days of future Quantitative Precipitation Forecast (QPF) grids. The NCRFC starts the FMAP process by populating the time-series with QPF from the NWS Weather Prediction Center (WPC). Then a comparison is made with QPF forecasts generated by the individual Weather Forecast Offices (WFOs). During heavy rainfall or snowfall events the WFOs might alter their forecasts from the national WPC guidance due to local knowledge, regional analysis, and further investigation into the weather situation.

3.3 Forecast Models Used for RRAF Generation

A description of the NCRFC modeling operations was introduced in section 1.4. Recall that there are two major models used in every basin and they are run on a 6-hour time-step. The two models, described in more detail in the following two sub-sections, are the Snow-17 and SAC-SMA models. Output from those models are used to generate the RRAF guidance. Operationally, the model runs supporting the RRAF are ran 3 times daily, currently around 0800L, 1100L, and 2200L. The two morning issuances are based on the 12Z model runs and the evening run is based off the 00Z model run. The 0800L guidance is thought of as an early look with minimal forecaster interaction whereas the 1100L run is considered a final run where forcing data has been thoroughly quality controlled and forecasters have evaluated the models and made modifications as necessary. Once the model runs are finished automated scripts transfer the lists of simulated events and other datasets over to DATCP for them to ingest into their database and populate the RRAF website.

3.3.1 Snow-17 Model

The first step of the model process in the NCRFC framework is running the Snow-17 Snow Accumulation and Ablation Model. It is a conceptual model which requires only temperatures and precipitation to simulate snow accumulation and snowmelt for all basins (Anderson, 2006). The Snow-17 uses a temperature index model to simulate the energy exchange into and within a snowpack instead of a more complicated energy balance model for several reasons. The first reason centers on the observed data available. Air temperature is easily measured and therefore measured at many locations in the U.S. whereas other energy exchange variables were not measured very often or available in real time when Snow-17 was developed (Anderson, 2006). Secondly, spatial variability of air temperature is fairly well understood and it is one of the easier variables meteorologists can predict. Most importantly, Anderson (1976) found a calibrated temperature index model produced similar output when compared to a more detailed energy balance model.

When developing the Snow-17 model, the goal was to first simulate the physical processes that occur in a column of snow. Further improvements allowed the model to be applied on an areal basis. To do this the model has twelve parameters that can be set during calibration and the model simulates 14 state variables during each time-step. The end result of the Snow-17 model is the total outflow from the snow pack and rain on bare ground, referred to as Rain + Melt (RAIM), which is fed as input into the hydrologic model. RAIM, in depth over the basin (mm), is defined as the following:

$$RAIM = (O_s \cdot A_s) + [(1.0 - A_s) \cdot (P \cdot f_r)]$$

Where the definitions are as follows:

O_s = outflow from snow cover (mm)

A_s = Areal extent of snow cover (decimal fraction)

P = Total precipitation (mm)

f_r = Fraction of precipitation in the form of rain (decimal fraction)

For a relatively simple model, NCRFC forecasters have noted that the Snow-17 has consistently provided decent results over the years. Handling rain on snow events and the capability of seasonally adjusted melt rates are examples of why this temperature index approach performs so well. Operationally, forecasters have the ability to modify the Snow-17 model if the need arises. As observers report snow water equivalent measurements and airborne surveys are completed by the National Operational Hydrologic Remote Sensing Center (NOHRSC), forecasters have the ability to compare their simulated snow states versus the real world measurements. Adjustments can be made by removing or adding moisture to the snow pack. Another common modification allows forecasters to speed up or delay the snowpack melt rate to match river gauge response. For a more detailed review of the Snow-17 model and how it simulates areal snow cover and outflow on a time-step basis reference the report written by Anderson (2006).

3.3.2 Sacramento Soil Moisture Accounting Model

The rainfall plus snowmelt time-series from the Snow-17 model is subsequently passed to the SAC-SMA model for each basin to compute runoff and eventually streamflow. The SAC-SMA model was developed at the California-Nevada River Forecast Center (CNRFC) in Sacramento, California in the early 1970s after it was determined the Antecedent Index model routinely produced poor streamflow simulations in the western U.S. It was determined a more complex model was needed to simulate the physical processes involving runoff on a watershed scale. However, both limited computational capabilities and limited observed data led the developers to choose a middle ground solution between complex and simple models. This decision led to both improved streamflow modeling as well as allowing the forecast model to be dependable and used in operational, day-to-day forecasting. The resulting SAC-SMA model is a conceptual, lumped, continuous two-layer soil moisture accounting model which requires only precipitation (and/or snowmelt) and an evapotranspiration curve as inputs. It is important to remember, especially with regard to the RRAF, that the SAC-SMA model was developed, and calibrated, with the goal of providing estimates of watershed discharge. In addition, due to the lumped nature of the model, only one value per variable is valid for the entire basin at a given time-step, even if the basin is several hundred square miles in size.

The following paragraphs will briefly discuss some of the SAC-SMA parameters with an emphasis on the upper zone components that are used in the RRAF development. For a more detailed review of the SAC-SMA model a good reference is written by Burnash (1995). As mentioned earlier, the SAC-SMA model contains two conceptual soil zones containing moisture storage areas. In the upper soil layer the storage areas are referred to as the Upper Zone Tension Water (UZTW) and the Upper Zone Free Water (UZFW). Tension water refers to moisture in the soil that is held tightly in place with the pull of gravity balanced by the molecular attraction to soil particles and water molecules. In general, the potential volume of tension water is determined by interrelationships of a basin's climate, soil types, and vegetation coverage and variety. Tension water can only be extracted from the soil via evapotranspiration (ET) (Burnash, 1995).

The UZFW represents liquid water residing in the void between soil particles that is free to percolate deeper into the ground or move laterally through the soil. In physical terms, all rainfall is applied to the surface as "free water". What happens next is dependent on the dryness of the soil and the soil permeability. Deficiencies in the UZTW will result in moisture converted to tension water. After those requirements are met, this moisture is able to either percolate or move laterally towards a stream channel. The SAC-SMA algorithm follows this process closely. The first step is to fill UZTW needs. Remaining water is sent to the UZFW storage where percolation is accomplished. The magnitude of percolation varies with the lower zone moisture deficiencies with more water percolated when the lower zones are dry and less when they are saturated. Water remaining after tension water deficiencies and percolation are met is allowed to discharge laterally to the channel (Burnash, 1995).

The SAC-SMA produces runoff in 5 different ways and it should be noted that the runoff component names are not completely analogous with common hydrologic terminology. The first three runoff components occur in the upper zone while the last two are components of base flow and will not be discussed here. The runoff components are: Direct Runoff, Interflow Runoff, Surface Runoff, Supplemental Base Flow Runoff, and Primary Base Flow Runoff. Runoff, like most parameters for the SAC-SMA model, is expressed as a depth over the basin.

The SAC-SMA model allows two types of surfaces in a basin, permeable and impervious. A unique feature of the SAC-SMA is that the percent impervious area in a basin does not have to remain constant. Additional impervious area can be activated by increasing soil moisture in the upper zone thus mimicking the act of filling marshes or small detention basins that are linked to the stream channel. When incoming rainfall and snowmelt are applied to the basin, the areal percentage of the basin deemed as permanent or temporarily impervious produces Direct Runoff. Not every basin has impervious area defined, and when it is defined, it generally is not a very large percentage. Therefore, Direct Runoff is not universally simulated across an area such as Wisconsin nor is it a significant percentage of the runoff volume when produced.

The second upper zone runoff component is termed Interflow Runoff (INTRO). This runoff is generated by the process described in the prior UZFW discussion. When RAIM is applied to the permeable portion of a basin, the UZTW is filled until completely full. Excess moisture then goes to UZFW and percolation begins. After percolation demand is met, excess UZFW is allowed to drain to the channel as the conceptual lateral flow called Interflow Runoff. The drainage rate from UZFW is defined during calibration using a depletion coefficient called UZK. Therefore for each time-step:

$$INTRO = UZFW\ Contents\ \cdot UZK$$

Generally, INTRO is a large percentage of a basin's quick response runoff as well as being fairly common in simulated basin runoff behavior across the NCRFC region. Note that some basins can have a delayed response if an unusually small UZK coefficient was defined during calibration.

Surface Runoff is the third upper zone component and by definition occurs generally during heavy rainfall events. In order to generate Surface Runoff in the SAC-SMA, the incoming RAIM rate must fulfill UZTW deficiencies, exceed percolation rate and INTRO drainage rate, and fill the UZFW completely. Excess water after those demands are met will be discharged as Surface Runoff to the channel. As this description suggests, generation of Surface Runoff is not very common and generally reserved for heavy rainfall or long lasting snowmelt events that exhibit extended high rates of water into the system.

3.4 Definition of a Simulated Runoff Event

A lengthy analysis of model output versus observed runoff events determined the definition of a simulated runoff event to be used for the RRAF product. A modeled event is defined by three model components meeting the following required conditions for a given time-step:

1. Snow-17: Rain + Melt (RAIM) > 0
2. SAC-SMA: Interflow Runoff (INTRO) > 0
3. SAC-SMA: Upper Zone Tension Water Deficit (UZTWD) = 0

Where the Upper Zone Tension Water Deficit (UZTWD) is defined as:

$$UZTWD = UZTWM - UZTWC$$

The UZTWD represents the amount of water needed to completely fill the UZTW storage. The maximum volume of the UZTW storage, Upper Zone Tension Water Maximum (UZTWM), is set during calibration and can change from basin to basin. The amount of water in the UZTW at any given time is indicated by the Upper Zone Tension Water Contents (UZTWC) which is capable of increasing or decreasing every time-step during the simulation.

As section 3.3 suggested, INTRO was chosen as the main event indicator as it is the most reliable “surface-based” runoff component from the SAC-SMA. It is fairly fast responding in nature and simulated consistently across Wisconsin. However, early testing indicated that INTRO alone was not a sufficient approach for field scale runoff risk. Requiring RAIM to be greater than zero focused the event durations to center around the time precipitation is falling or snow is melting. This also eliminated simulated events of long duration (Interflow present for several days) due to small UZK coefficients and thus very low UZFW drainage rates in a basin. Similarly, the UZTWD component was also added to concentrate the simulated events on the risky conditions when upper soils (conceptually) are saturated and runoff at the surface is most likely to occur.

The beginning of a simulated event is designated as the first time-step the three conditions are met and the event ends with the first time-step at least one condition is no longer true. However, some post-processing is necessary before the final simulated event duration (beginning time and ending time) is determined due to the nature of weather and hydrologic models. These models, like most others, run on an accumulated time-step basis. For example, precipitation at 18Z is the accumulated total from the previous time-step up to that time. In this case the 18Z precipitation value is the amount generated from 12Z to 18Z. Due to this, the event beginning time is defined as the time-step all three conditions are met minus 6-hours to incorporate the inferred accumulated time.

The end time of the simulated event is also adjusted by adding a 6 hour time-step “buffer” to the time when at least one of the criteria is no longer met. This was done as a hedge to compensate for the possible delay in INTRO being generated from the lumped SAC-SMA model. For example, if one of the event conditions was false at

06Z the event end time would be adjusted to include one more time-step and be defined as 12Z.

For the event duration the INTRO is accumulated to provide a total event runoff depth. This event depth is the key indicator used for basin runoff risk for the RRAF. A runoff type category is also included with each event to indicate the cause of the runoff. These runoff type flags listed below are determined by comparing the accumulated RAIM value during the event with the accumulated liquid portion of the forecast precipitation designated by Future Mean Areal Precipitation (FMAP):

- F0: Event due to only rainfall
  - $When\ (RAIM = FMAP)$
- F1: Event due to combination of rain and snowmelt
  - $When\ (RAIM\ > FMAP)\ \&\ (FMAP\ > 0)$
- F2: Event is due to only snowmelt
  - $When\ (RAIM\ > 0)\ \&\ (FMAP\ = 0)$

3.5 Generation of Historical Simulated Runoff Events

To conduct the RRAF validation with observed runoff events it was necessary to generate a dataset of simulated runoff events. This was accomplished using the NCRFC Ensemble Streamflow Prediction (ESP) system. The ESP system is used operationally by the RFCs to generate probabilistic flow and stage forecasts by comparing a historical simulation against a "conditional simulation" derived from a distribution of forecast runs. For the RRAF analysis, the historical simulation was used and consisted of a continuous simulation from 1948 to 2008 (6-hour time-step). The historical model run included quality-controlled MAP and MAT forcing data based on observations which eliminate the forecast component. However, it should be remembered that the historical run is still a simulation and therefore will differ from observed streamflow. Using observed forcing data provides the best representation possible of past conditions by the model.

From the historical run, the three model components used for the RRAF were extracted and the simulated runoff event code was applied to produce a list of runoff events for every basin in the historical time frame. From the historical distributions of the NWS test basins, a subset of events was selected that correspond with the observed runoff time frames. This is to ensure a fair analysis can be produced without injecting invalid false alarms or event misses. Figures 8 and 9 highlight the historical simulations of the NWS test basins used in the validation. The dashed lines represent the entire historical distribution where the solid line represents the range of events found in the comparison.

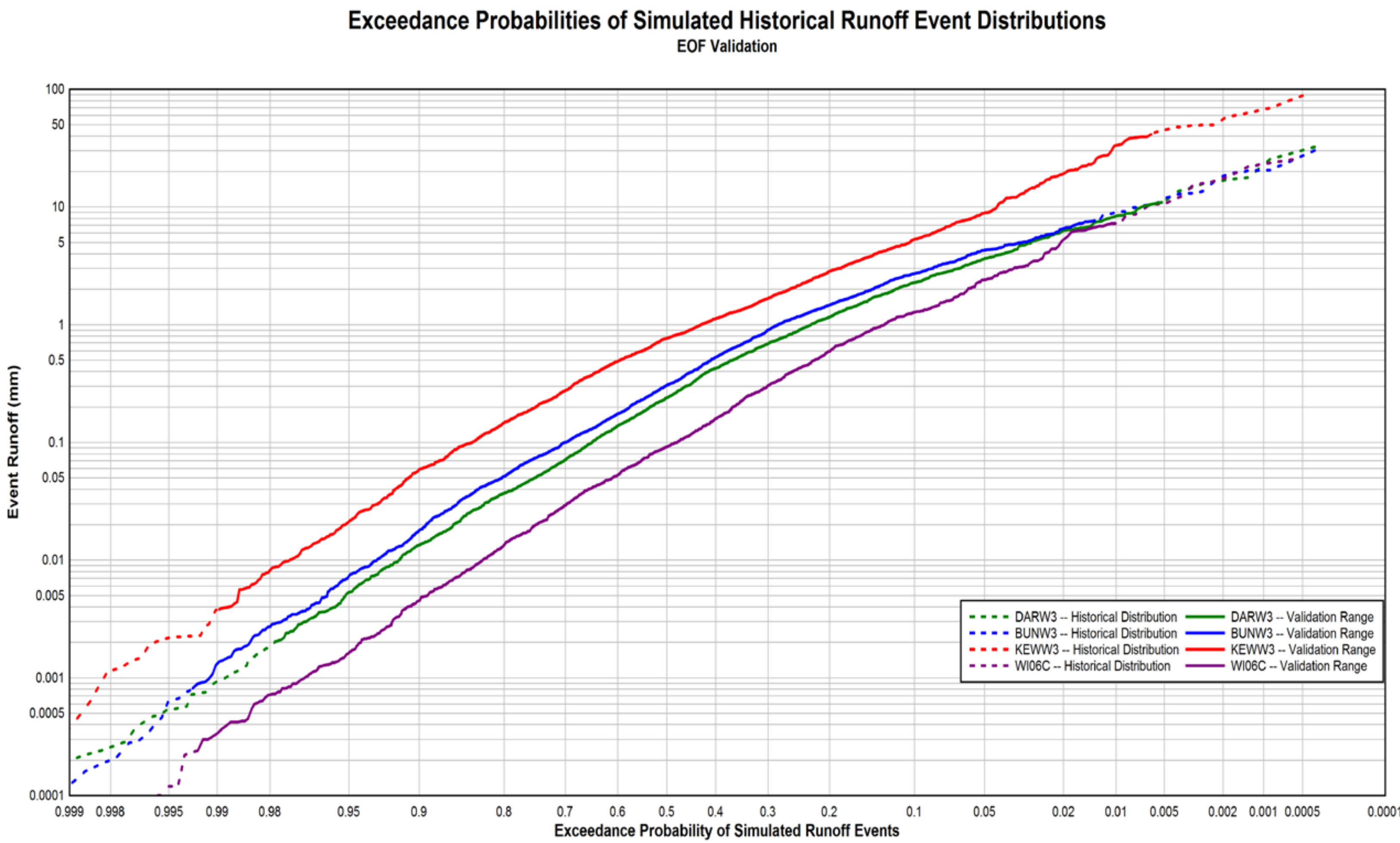


Figure 8. Exceedance probabilities of simulated historical runoff events for four NWS basins compared with observed edge-of-field runoff data. Dashed lines represent the entire distribution where the solid line represents the range of events included in the validation time frame.

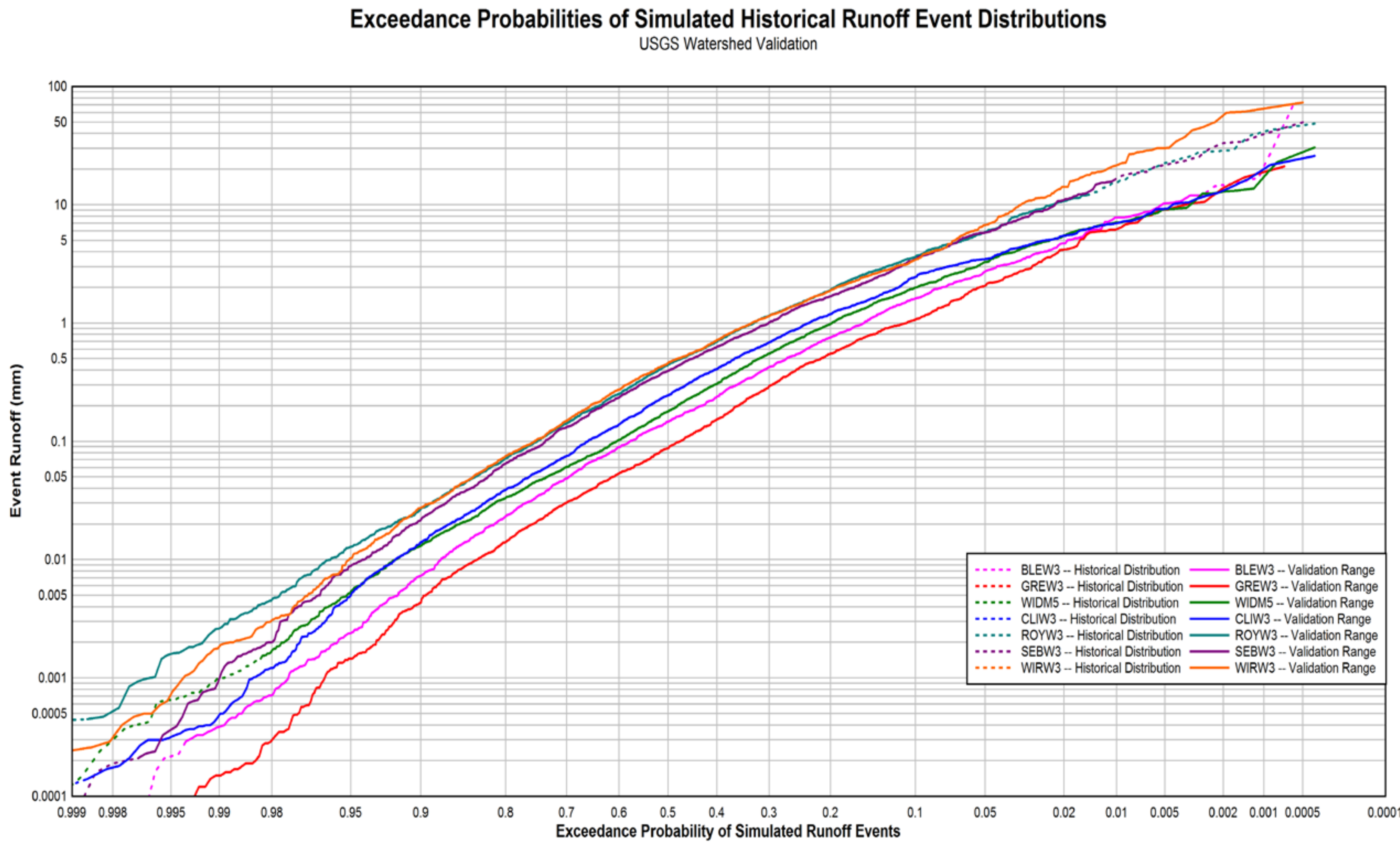


Figure 9. Exceedance probabilities of simulated historical runoff events for seven NWS basins compared with observed edge-of-field runoff data. Dashed lines represent the entire distribution where the solid line represents the range of events included in the validation time frame.

3.6 Summary of Observed Field Scale Runoff Data

The next two sections will review the observed historical data used in the validation of NCRFC historical model output to generate a runoff risk decision support tool. Observed runoff datasets were gathered for two spatial scales. The first observed dataset described is the edge-of-field (EOF) runoff data provided by the USGS and Discovery Farms. Runoff data was collected using H flumes installed in waterways (Figure 10) where runoff from a given field could be accurately measured and recorded. Reference Stuntebeck et al. (2008) for a detailed review of the challenges and procedures of collecting this data. Many variables were monitored at these test locations, however only the processed event runoff data was used in this study.

Four edge-of-field datasets from Discovery Farms (Figure 11) were used in the validation. Table 2 lists the areas of the monitored fields as well as the NWS basins they were located in. Comparing the average EOF area against that of the average NWS basin highlights the extraordinary challenge of the RRAF and the key focus of the validation: Can reasonable field scale runoff risk be derived from modeled watersheds of considerably larger size difference to aid in timing decisions of applying manure and nutrients? Specifically, with the EOF validation, the average area comparisons are between NWS basins over 145,000 acres in size and EOF basins of 21 acres. Another perspective can be thought of as the examination of the accuracy of the NWS watershed simulations based on observations representing only one-hundredth of a percent of that modeled basin area. A brief description of the four EOF sites follows, however a more in-depth summary is provided in Stuntebeck et al. (2011).

Figure 10. Example of USGS EOF measurement setup at a Discovery Farms site. For further description of equipment and measurement techniques see Stuntebeck et al. (2008).

Table 2. Areas of edge-of-field measurement sites and the corresponding NWS basin they are located in.

		mi²	km²	acres
County EOF	**NWS Basins**			
Iowa	DARW3	0.025	0.064	15.8
Lafayette	BUNW3	0.062	0.160	39.5
Manitowoc	WI06C	0.028	0.071	17.6
Kewaunee	KEWW3	0.021	0.053	13.2
EOF Average		*0.034*	*0.087*	*21.5*
NWS Basins	**EOF Basins**			
DARW3	Iowa	273	707	174,720
BUNW3	Lafayette	125	324	80,000
WI06C	Manitowoc	384	995	245,870
KEWW3	Kewaunee	127	329	81,280
NWS Average		*227*	*589*	*145,468*

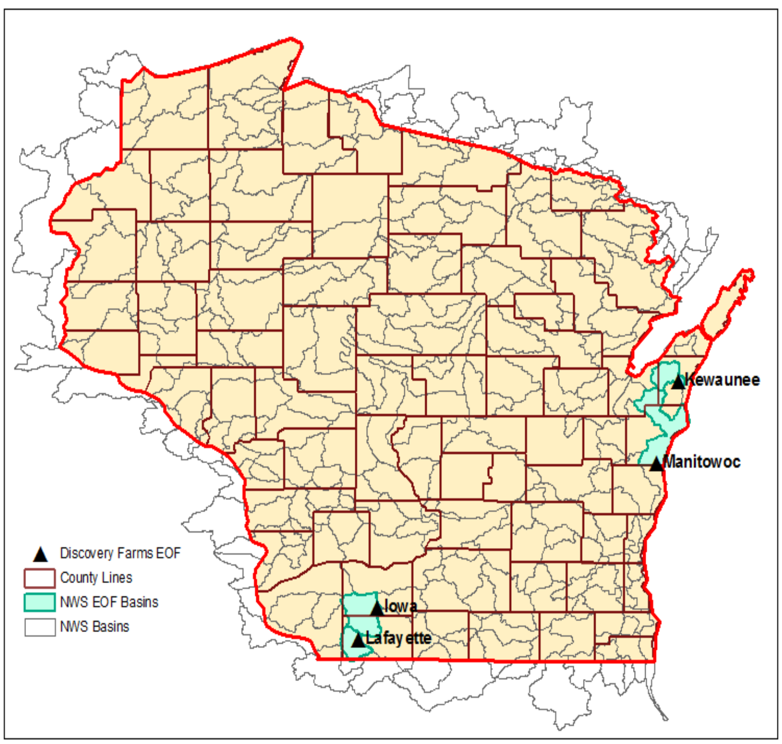


Figure 11. Location of Discovery Farms edge-of-field runoff monitoring sites used in RRAF validation.

3.6.1 Iowa County Edge-of-Field Runoff Summary

Located in Iowa County on the southern edge of the Driftless Area is the first observation site. The studied watershed is found on a dairy farm with 110 milk cows. Soils are a fine, silty loam characterized by low permeability and classified as Dodgeville silt loam (Stuntebeck et al. 2011). Nutrient management on the 320 acres of cropland was conducted with a phosphorus-based NMP where all cropland received one application of manure on a 4-year rotation at a rate of 15-20 ton/acre (Stuntebeck et al. 2011). The EOF site (Figure 12) was 15.8 acres with the flume measuring runoff located in a grassed waterway. The watershed contained all or parts of seven separate fields which were always planted with either corn or alfalfa. Slopes in the watershed ranged from two to eight percent (Stuntebeck et al. 2011). Runoff was monitored at this site from July 2004 through March 2007. A summary of the runoff events can be seen in Table 3.

Table 3. Summary of observed runoff events for Iowa County EOF site.

Iowa County EOF Event Runoff (in)			
Number of Events	36	Observations	Jul 2004 – Mar 2007
Average	0.1306	Minimum	0.0008
Median	0.0494	Maximum	0.8724
Standard Deviation	0.1950		

3.6.2 Lafayette County Edge-of-Field Runoff Summary

The second EOF site is located on a 600-head beef-finishing farm in Lafayette County. The cropland on this farm has not been tilled in 20 years and consists of a fine, well-drained silty loam named Tama (Stuntebeck et al. 2011). The test watershed was 39.5 acres (labeled R3 in Figure 13) and consisted of two fields. The total farm size is near 800 acres of cropland in mostly corn and soybeans. Manure harvested from the feedlots is routinely spread in the fields in January to April and October to December. Manure generated in the summer was stored and applied in the fall at a rate of 13 to 15 ton/acre (Stuntebeck et al. 2011). The monitoring period at this location lasted from February 2004 to August 2007 and a summary of runoff events is shown in Table 4.

Table 4. Summary of observed runoff events for Lafayette County EOF site.

Lafayette County EOF Event Runoff (in)			
Number of Events	65	Observations	Feb 2004 – Aug 2007
Average	0.0847	Minimum	0.0001
Median	0.0215	Maximum	1.0427
Standard Deviation	0.1677		

Figure 12. Aerial overview of the 15.8 acre edge-of-field measurement site in Iowa County. Figure is provided by Stuntebeck et al. 2011.

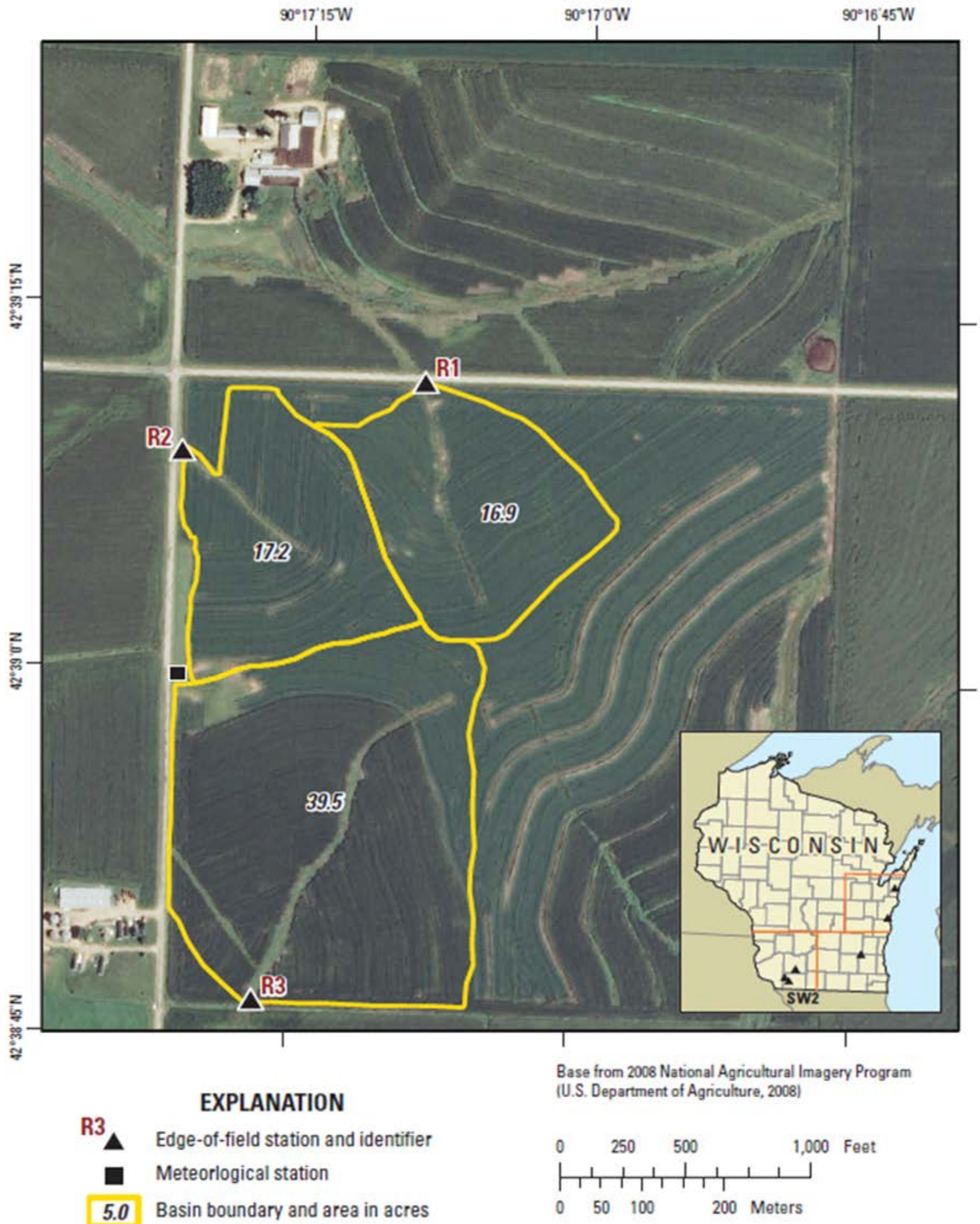


Figure 13. Aerial overview of the 39.5 acre edge-of-field measurement site (R3) in Lafayette County. Figure is provided by Stuntebeck et al. 2011.

3.6.3 Manitowoc County Edge-of-Field Runoff Summary

The EOF site in Manitowoc County was another dairy operation with a herd of over 1,000 animals. The 925 acres of land on this farm were rotated among grazing pasture, corn, and alfalfa. The operation followed a phosphorus based NMP and the soils were characterized as Kewaunee loam which is high in clay and thus has low permeability (Stuntebeck et al. 2011). As the herd was grazing most of the year, 50% of the manure was spread naturally while the other half was collected during winter housing and spread as LDM later. The observation site at this farm (labeled K3 in Figure 14) was active from December 2004 to April 2007 with the summary of runoff events in Table 5. The watershed monitored was 17.6 acres in corn/soybean rotation with slopes between 2 and 6 percent (Stuntebeck et al. 2011).

Table 5. Summary of observed runoff events for Manitowoc County EOF site.

Manitowoc County EOF Event Runoff (in)			
Number of Events	62	Observations	Dec 2004 – Apr 2007
Average	0.2178	Minimum	0.0001
Median	0.0357	Maximum	1.3345
Standard Deviation	0.3628		

3.6.4 Kewaunee County Edge-of-Field Runoff Summary

The last observed runoff site was located on a CAFO with 1,400 dairy cows in Kewaunee County. Mostly corn, alfalfa, and wheat were grown on this 1,800 acre farm which resides on low permeable soils classified as Hortonville silt loam (Stuntebeck et al. 2011). Manure was applied to fields according to a phosphorus-based NMP. Solid manure was generally applied 3 times a year at 5 ton/acre rate while LDM was applied at a rate of 4,000 to 30,000 gal/acre after alfalfa harvest in the summer and after corn harvest in the fall (Stuntebeck et al. 2011). Runoff was observed from November 2003 to March 2007 on a 13.2 acre single-crop watershed (labeled P3 in Figure 15) with slopes between 2 and 6 percent. This test field also had subsurface tile to help drain the wet clayey soils. Table 6 summarizes the runoff data from this site.

Table 6. Summary of observed runoff events for Kewaunee County EOF site.

Kewaunee County EOF Event Runoff (in)			
Number of Events	95	Observations	Nov 2003 – Mar 2007
Average	0.1862	Minimum	0.0002
Median	0.0499	Maximum	2.1588
Standard Deviation	0.3586		

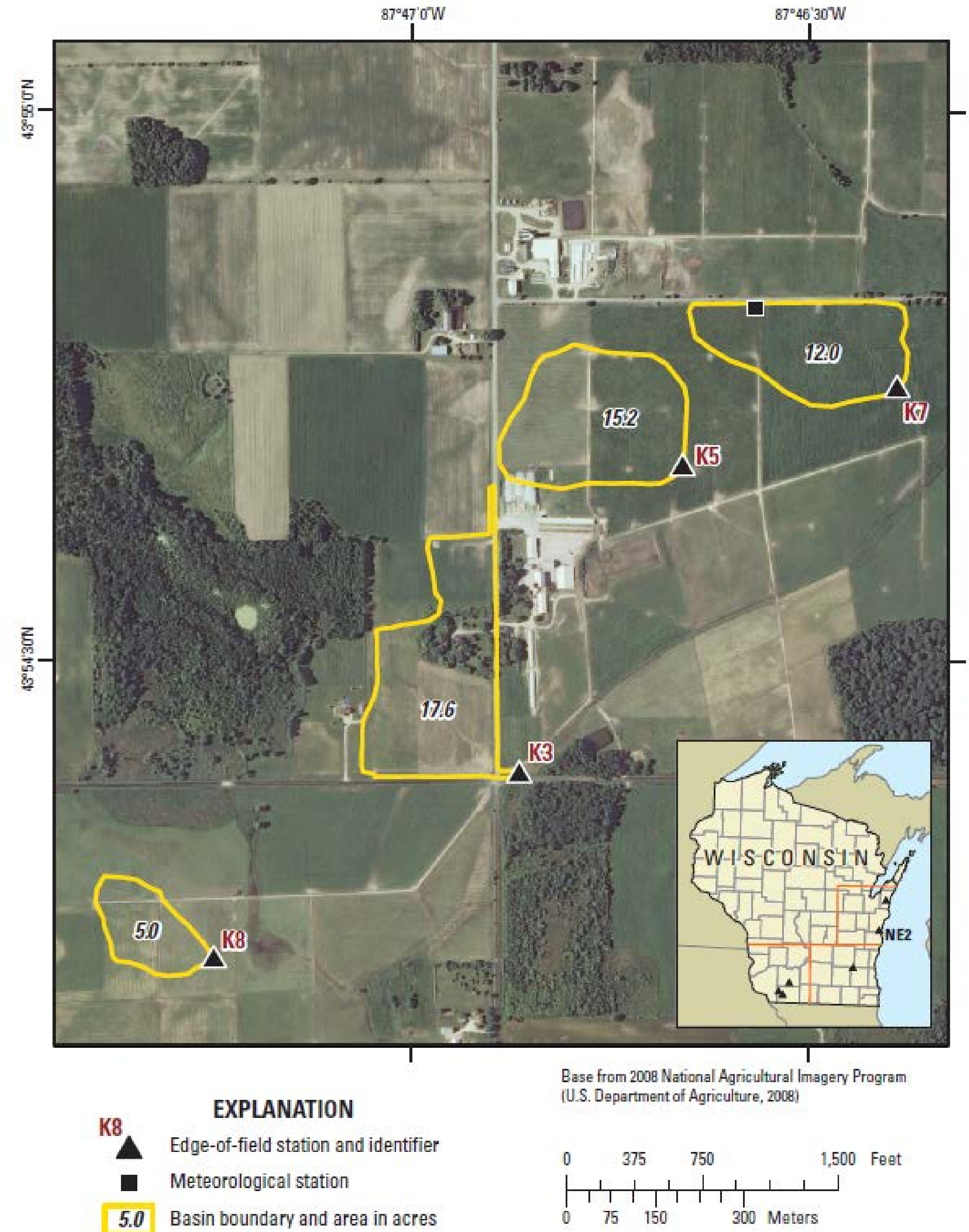


Figure 14. Aerial overview of the edge-of-field measurement site (K3) in Manitowoc County. Figure is provided by Stuntebeck et al. 2011.

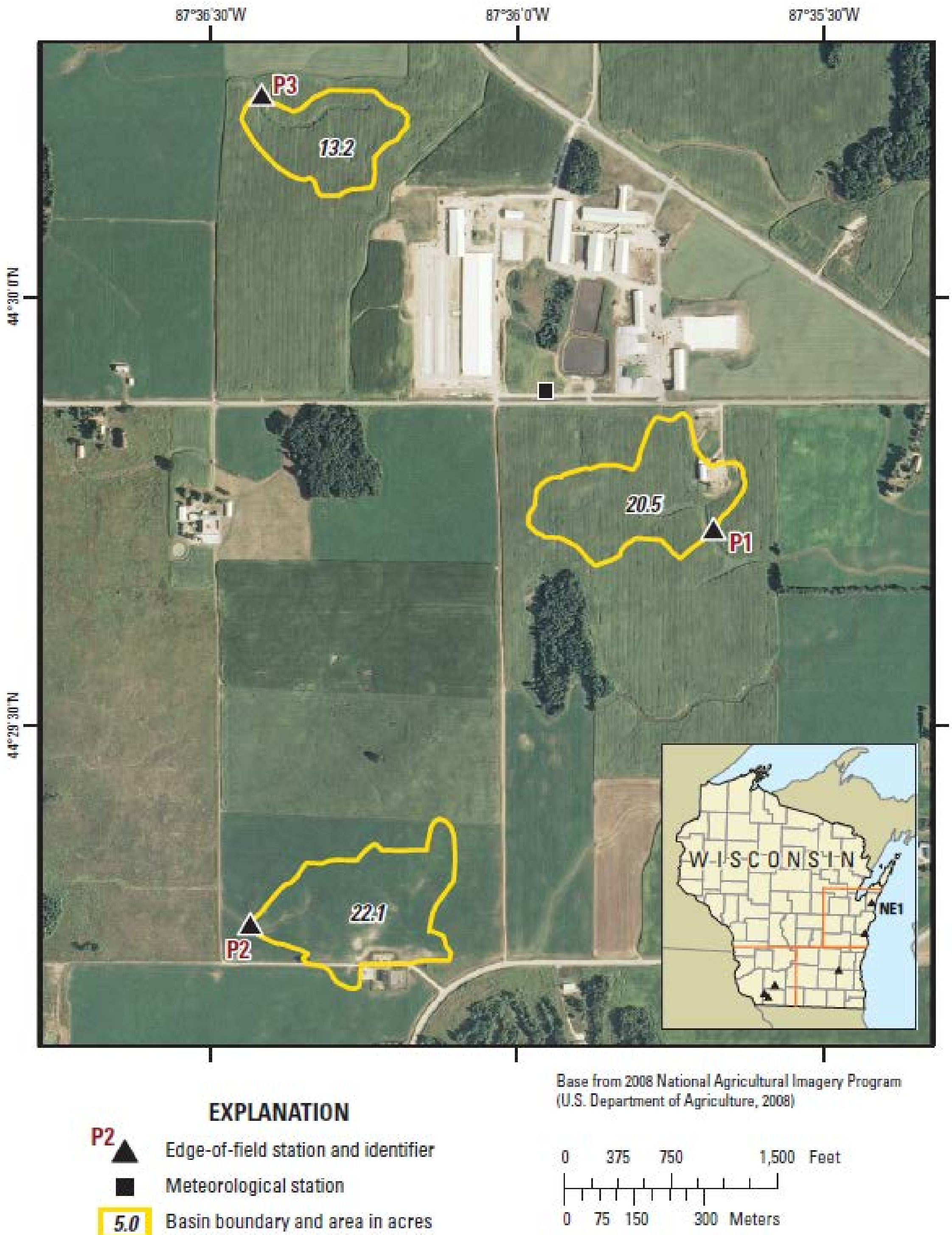


Figure 15. Aerial overview of the edge-of-field measurement site (P3) in Kewaunee County. Figure is provided by Stuntebeck et al. 2011.

3.6.5 Justification for using Small Watershed and Field Scale Runoff Data

The obvious scale difference between the modeled watersheds and EOF sites was recognized early on in the RRAF development process. The approach of comparing runoff response from an EOF site to larger NWS basins, where the EOF size is equivalent to ~0.01% of the average NWS basin size, does invite potential for misleading or inaccurate validation results. However, this analysis is still important as it evaluates the true viability of the RRAF to be useful for the decision makers on the field scale.

The EOF validation weaknesses center more on the lack of ability to describe how well or how poorly the RRAF model works. This is mostly a factor of the scale issue and the inability to confidently rely on the false alarm and miss statistics that will be generated. False alarms are when a model simulates an event and none is observed. A miss is defined as when an observed event occurs but none is simulated. A more in-depth discussion on false alarms and misses will be discussed starting in section 3.8.

It is easily seen how the scale dilemma can be used to cast doubt on the RRAF effectiveness using only one EOF site in a considerably larger NWS basin. The possible questions are many, however the answers are not easily known. Does that EOF site represent the average basin conditions? Is it skewed one way or the other in terms of runoff response? What is the spatial variability of precipitation like in that test basin? Therefore, it should be easy to see that it is plausible that EOF validation could induce a bias of too many false alarms or too many misses.

An example of a snowmelt scenario could be used to demonstrate how EOF sites could bias the RRAF effectiveness analysis. Take a field with a southern aspect in the winter. Stronger solar radiation could help this field lose its snowpack earlier than the average field conditions in this NWS basin. If a snowmelt runoff event is then simulated, this field, having no actual snow, may not observe a runoff event and thus indicate a false alarm. The opposite scenario is also plausible where a northern facing field holds its snowpack longer and a runoff event is observed when the simulation shows no event and thus producing a miss. An accumulation of instances such as these over time would bias the analysis for this EOF validation.

If the validation of an EOF did produce a large value of false alarms or misses, it would be extremely challenging to prove or disprove that the EOF site itself is causing the bias compared to the rest of the area inside that NWS basin. Therefore, poor validation statistics could dampen the perceived RRAF effectiveness when perhaps it actually does a very good job on average across that basin.

To get around the EOF-only induced limitations, one of the two following options would be ideal: (1) have a dense network of EOF sites included in a NWS basin, or (2) have much smaller NWS basins where one EOF observation site represents a significant portion of that modeled area. As with many scientific investigations where perfect datasets do not exist, the two ideal recommendations are not plausible at this time. The dense network of EOF sites would require an extensive budget to accomplish. The second suggestion is invalid by its definition, as the RRAF development goal was tied to using “current” NCRFC modeling capabilities. However, in the future, progressing to a finer resolution distributed model could be possible and thus allow a more justifiable analysis with only one EOF dataset per modeled basin.

Given the limitations of the one-EOF-per-basin validation, it was decided to include another runoff dataset to validate against the model. The use of derived runoff events from small watersheds gauged by the USGS was selected. It is assumed that these small watersheds would more appropriately mirror NWS basins by including a mixture of physical landscape attributes. Their larger size essentially mimics having many EOF sites bundled together to be evaluated against the much larger NWS basins. There are two key assumptions being made with the inclusion of this data to the validation: (1) simulated false alarms and misses should decrease due to larger observed area being tested, and (2) the validity of the observed runoff events remains applicable to the farm field scale as the watersheds are still fairly small in size.

3.7 Summary of Observed Small Watershed Runoff Data

Several guidelines were initially defined to select a sample of small watersheds for the validation. A survey of mean daily flow data available from the USGS in Wisconsin was completed using the following proposed conditions:

- At least 5 years of mean daily flow in the last 20 years.
- Stations representing a variety of Wisconsin geomorphic conditions
- Watersheds near the size of 10 - 25 mi^2

Although it was not possible to meet all of these guidelines, seven small watersheds (Figure 16) were selected to validate against the model data. Once the mean daily flow data was obtained the next step required extracting the runoff from the total streamflow and then defining when these events begin and end. The work to produce a list of runoff events, with starting and ending times as well as runoff as a depth over the watershed, for each of the seven small watersheds was performed by Seth McClure (Wisconsin DATCP) and Dr. John Panuska (University of Wisconsin – Madison Extension) (McClure, 2009).

To generate the runoff events the U.S. Bureau of Reclamation's Base Flow Index (BFI) software was used (Wahl & Wahl, 2007). The Base Flow Index method estimates the annual base flow volume for unregulated streams and produces an annual index consisting of the ratio of base flow to total flow. Over the years this method has been shown to be capable of producing a base flow time-series equivalent to more sophisticated approaches (Wahl & Wahl, 2007).

The next step for McClure and Panuska was to decide how to define runoff events from the BFI data. Recommendations from Dave Graczyk (Hydrologist, USGS Middleton Wisconsin) were implemented and consisted of the following:

(1) A separation of at least six hours between storm flow peaks would be used to separate events.
(2) An 85% BFI threshold was used to extract storm runoff from base flow.
(3) The duration storm flow runoff was allowed to continue was based on a defined lag time.

The 85% BFI threshold rule is used to define when storm runoff existed. Essentially if the calculated base flow was greater than 85% of the total flow for that day there was no storm runoff. For the third condition, it was decided that a cut off point for storm runoff was needed and would be proportional to the watershed size. This "lag time" was defined as:

$$t_{lag} = DA^{0.2}$$

Where t_{lag} is the time in days a storm runoff event is allowed to continue and DA is the drainage area (in mi^2) of the watershed. For all seven watersheds selected the lag time was found to be 2 days. To summarize, when storm flow is detected (85% of the total flow > than calculated base flow) any flow for the following two days above base flow is considered to still be storm runoff and part of its hydrograph (McClure, 2009).

The selected USGS watersheds on average were about 15.9 mi^2 (Table 7), up considerably from the 0.03 mi^2 of the EOF sites. The NWS basins corresponding to the USGS watersheds averaged 294 mi^2, also a bit larger than the average value for the EOF comparison (227 mi^2). This portion of the average NWS basin area being "evaluated" by the observed data increased to 5.4% from 0.015% for the EOF sites. The time span included in the USGS watershed validation is generally longer (Table 8) than what is used for the EOF sites with more observed events to help with the comparison.

Table 7. Summary of USGS watersheds and corresponding NWS basin areas.

			mi^2	km^2	acres
USGS Basins	**USGS ID**	**NWS Basin**			
Otter Creek	40857005	SEBW3	9.5	24.6	6,080
Eagle Creek	05378185	WIDM5	14.3	37.0	9,152
Black Earth Creek	05406460	BLEW3	14.6	37.8	9,344
Bower Creek	04085119	GREW3	14.8	38.3	9,472
Jackson Creek	05431016	CLIW3	16.8	43.5	10,752
Little Plover River	05400650	WIRW3	19.0	49.2	12,160
Little Wolf River	04079602	ROYW3	22.6	58.5	14,464
USGS Average			*15.9*	*41.3*	*10,203*
NWS Basins					
SEBW3			425.1	1,100.9	272,038
WIDM5			318.1	823.8	203,565
BLEW3			46.4	120.3	29,727
GREW3			124.1	321.4	79,420
CLIW3			191.2	495.2	122,367
WIRW3			455.1	1,178.7	291,263
ROYW3			498.3	1,290.7	318,939
NWS Average			*294.0*	*761.6*	*188,188*

Table 8. Summary of BFI derived runoff events (in) from chosen small USGS gauged watersheds used to compare against simulation data.

	Otter Creek	**Eagle Creek**	**Black Earth Creek**	**Bower Creek**	**Jackson Creek**	**L Plover River**	**L Wolf River**
Start Obs	1 - 1991	1 - 1991	1 - 1990	1 - 1992	1 - 1994	1 - 1960	1 - 1974
End Obs	9 - 2002	9 - 2007	7 - 2005	7 - 2008	8 - 2008	9 - 1978	11 -1978
Num Events	228	191	75	110	273	229	84
Event Avg	0.2523	0.0851	0.1022	0.4154	0.3350	0.0422	0.1831
Event Med	0.0572	0.0243	0.0341	0.0278	0.0882	0.0147	0.0507
Event Max	4.7843	0.6770	1.1987	3.2034	5.2801	0.4704	2.9564
Event Min	0.0026	0.0031	0.0024	0.0000	0.0004	0.0028	0.0026

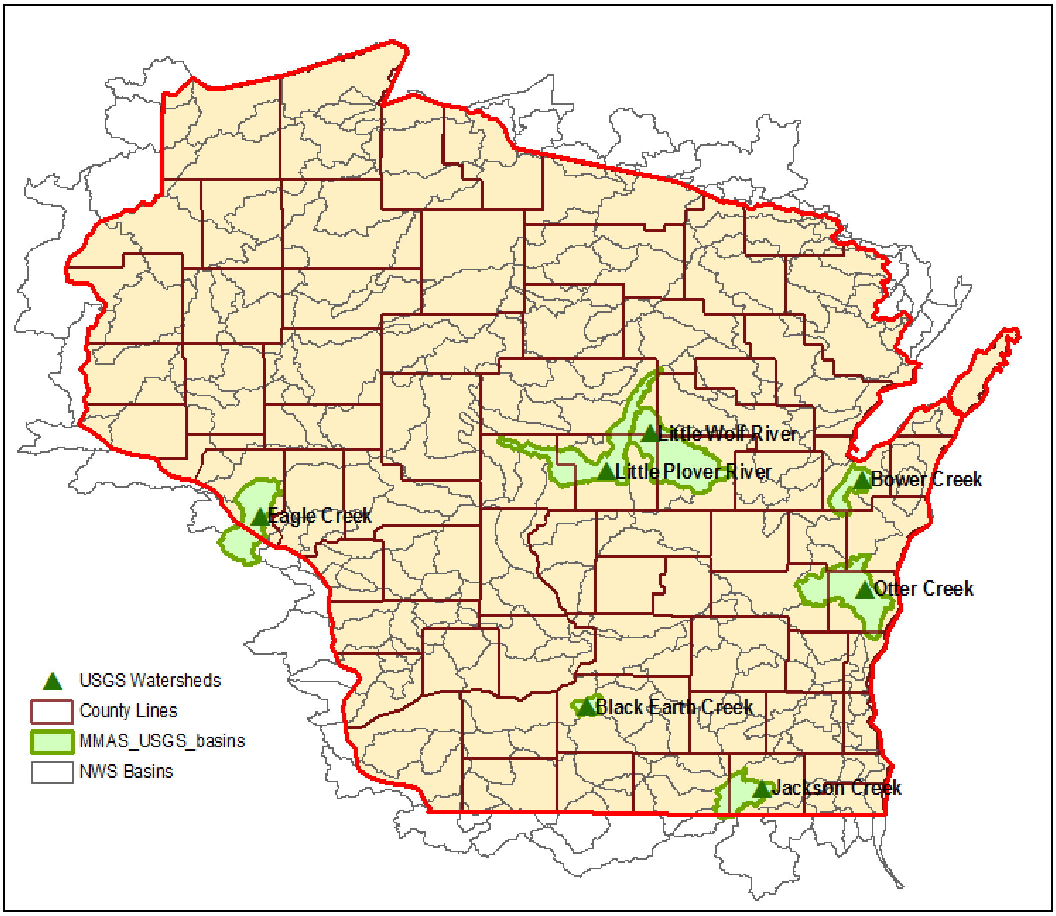


Figure 16. Location of the seven selected USGS gauged small watersheds used for RRAF validation.

3.8 Validation of Runoff Event Model vs. Observed Runoff Data

The goal of the validation of the simulated runoff events is simply to determine if the RRAF model can replicate the occurrence of observed runoff at the two spatial scales chosen. Further, if the model is viable, how well does it work? Does it over-alarm, or does it miss too many events?

Several metrics were chosen to help evaluate the RRAF effectiveness. Before those are discussed it is necessary to describe the validation technique chosen. Due to the challenges of spatial scale and the reality of a 6-hour model time-step, it was decided to use a generally simple comparison test. Python code was written to test the list of the observed events against the list of simulated events. Specifically the event start and end times were compared and any amount of overlap was identified. A simple schematic is shown in Figure 17. In the example the green colored events represent either observed (OE) or simulated events (SE) categorized as hits due to any overlap. The red shaded observed event is considered missed while the blue shaded simulated events represent false alarms.

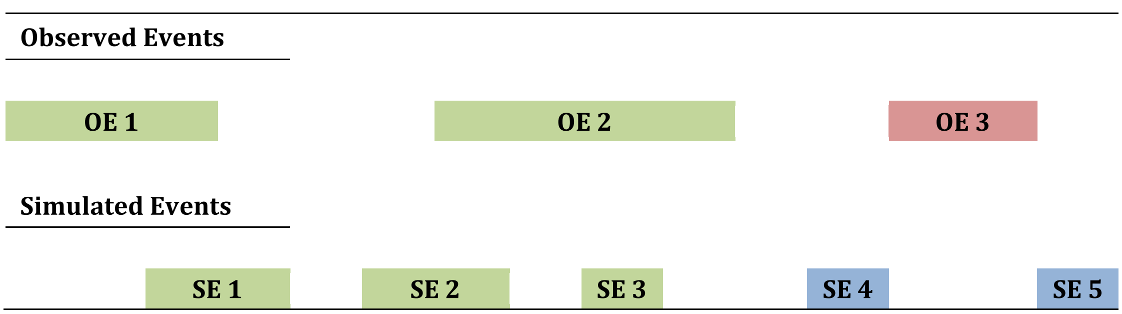


Figure 17. Schematic of the validation and verification method used to categorize a series of observed and simulated runoff events as hits (green), misses (red), or false alarms (blue). Numbers indicate the tally for either observed or simulated events.

The term "hit" could be thought of in two ways. An observed hit (H_o) was anytime an observed event had been recorded and there was a simulated event overlapping at least a portion of the observed duration. Similarly, a simulated hit (H_s) occurred anytime the model produced an event and there was a corresponding observed event overlapping at least part of that duration. Note that one event (simulated or observed) could verify one or more of the opposite type. For example, a long duration observed event could verify two or more simulated events that exist entirely during the observed event, and vice versa. The miss category (M) applies only to the observed runoff events. This occurs anytime a runoff event is observed yet no simulated runoff event overlaps it. The false alarm category (F) only applies to the simulated event dataset. This occurs anytime a simulated event is not overlapped by an observed event.

The comparison resulted in tallies of hits, misses, and false alarms for each of the test basins. These values can then be used to create summary metrics to describe how well the RRAF model captured the runoff risk. Four basic metrics were used and defined below: Bias, Probability of Detection (POD), False Alarm Rate (FAR), and Critical Success Index (CSI). A perfect score for Bias, POD, and CSI is a value of 1.0, where for FAR it is 0.0.

$$Bias = \frac{(H_O+F)}{(H_O+M)} \qquad CSI = \frac{H_O}{(H_O+M+F)}$$

$$POD = \frac{H_O}{(H_O+M)} \qquad FAR = \frac{F}{(H_S+F)}$$

As noted in section 3.6, observed runoff events were reported in inches. The simulated runoff events were created in millimeters. Remember the focus of the RRAF is not in producing the amount of runoff expected from a farm field. Rather the purpose is to focus on the likelihood of runoff occurring from fields in the NWS basin. Runoff depths for the separate datasets were not used to compare with each other, but instead used to stratify the response of that particular basin.

3.9 Field Scale Runoff Event Comparison

The four EOF observation site comparisons will be listed below followed by a summary of the combined EOF dataset after that.

3.9.1 Iowa County Edge-of-Field Runoff Event Comparison

The EOF site (15.8 acres) in Iowa County falls in the northeastern corner of the NWS basin DARW3 (174,720 acres) which is named after the Pecatonica River near Darlington, Wisconsin. The EOF site makes up only 0.009% of the total NWS area. For this comparison there were 36 observed events with 25 being classified as hits and 11 as misses. For the modeled events, there were 97 total with 17 classified as hits and 80 as false alarms. The summary metrics (Table 9) produced a value of 2.92 for the Bias and 0.22 for the CSI. The POD was a decent 0.69 while the FAR was 0.82. It is encouraging to note that 69% of the observed events were captured by the RRAF model. However, the high false alarm rate is a bit alarming, even with the perspective of the significant scale difference.

Looking closer at the observed runoff dataset a pleasant distinction exists between the values associated with hits and those categorized as misses. For example, the average and median event runoff value for hits is 0.1742 and 0.1026 inches. The average and median for the missed events are 0.0314 and 0.0166 inches respectively. That results in hit runoff depths being about 5 to 6 times larger than the missed events. The same type of stratification is seen in the event durations. The hit durations in hours and minutes are 15:34 and 13:26 (average and median) where the missed events span 4:09 and 0:34 (average and median) in length. The ratio of the hits to misses for the average duration values suggests the hits are 3.75 times longer than misses, however in comparing the medians, the difference is more dramatic with the hits being 23.7 times longer.

Table 9. Comparison of runoff events between Iowa County EOF and NWS DARW3.

Comparison of Runoff Events between Iowa County EOF and DARW3							
	Events	**Hit**	**Miss**	**FA**	**% Hit**	**% Miss**	**% FA**
OBS	36	25	11	--	69%	31%	--
SIM	97	17	--	80	18%	--	82%

Observed Event Runoff (in)		**Simulated Event Runoff (mm)**		
Hit Average	0.1742	Hit Average	1.6415	**Bias**
Hit Median	0.1026	Hit Median	0.7360	2.92
Miss Average	0.0314	FA Average	0.4806	
Miss Median	0.0166	FA Median	0.2706	**POD**
Avg Hit/Miss	5.55	Avg Hit/FA	3.42	0.69
Med Hit/Miss	6.18	Med Hit/FA	2.72	

Observed Event Duration (h:m)		**Simulated Event Duration (h:m)**		
Hit Average	15:34	Hit Average	54:00	**FAR**
Hit Median	13:26	Hit Median	24:00	0.82
Miss Average	4:09	FA Average	20:42	
Miss Median	0:34	FA Median	18:00	**CSI**
Avg Hit/Miss	3.75	Avg Hit/FA	2.61	0.22
Med Hit/Miss	23.7	Med Hit/FA	1.33	

Though the distinction was still present, the stratification between the simulated runoff event hit and false alarm sub-sets was smaller than that seen in the observed dataset. Shown in Table 9 the simulated Hit/FA ratios were between 2.7 and 3.4 (median and average) for the event runoff while the ratios for the event durations were smaller ranging between 1.3 and 2.6 (median and average).

3.9.2 Lafayette County Edge-of-Field Runoff Event Comparison

Located in the northeastern sector of the NWS basin BUNW3 (Galena River near Buncombe, Wisconsin) is the Lafayette County EOF site. As seen in Table 2, the 39.5 acre EOF site represents only 0.05% of the 80,000 acre BUNW3 watershed. In this comparison there were 65 observed events with 49 being hit and only 16 missed producing a 75% hit rate. On the simulated side there were 169 events simulated with only 34 being hits and 135 categorized as false alarms. That produced only a 20% hit rate and an 80% false alarm rate. Summary metrics (Table 10) found the Bias score to be 2.83 while the CSI was 0.25. The POD was 0.75 and FAR was 0.80. Lafayette County is very similar to Iowa County with regards to the simulated event statistics and slightly better than the already pleasing observed event statistics.

Table 10. Comparison of runoff events between Lafayette County EOF and NWS BUNW3.

Comparison of Runoff Events between Lafayette County EOF and BUNW3

	Events	Hit	Miss	FA	% Hit	% Miss	% FA
OBS	65	49	16	--	75%	25%	--
SIM	169	34	--	135	20%	--	80%

Observed Event Runoff (in)		**Simulated Event Runoff (mm)**		
Hit Average	0.1075	Hit Average	1.9885	**Bias**
Hit Median	0.0446	Hit Median	2.4786	2.83
Miss Average	0.0148	FA Average	0.5676	
Miss Median	0.0119	FA Median	0.1373	**POD**
Avg Hit/Miss	7.26	Avg Hit/FA	3.50	0.75
Med Hit/Miss	3.75	Med Hit/FA	18.05	

Observed Event Duration (h:m)		**Simulated Event Duration (h:m)**		
Hit Average	13:00	Hit Average	36:10	**FAR**
Hit Median	9:00	Hit Median	24:00	0.80
Miss Average	7:54	FA Average	19:49	
Miss Median	8:27	FA Median	12:00	**CSI**
Avg Hit/Miss	1.65	Avg Hit/FA	1.83	0.25
Med Hit/Miss	1.06	Med Hit/FA	2.00	

Resembling Iowa County, the Lafayette County observed runoff events also show a separation between the hits and misses in runoff magnitude with ratios indicating hits are between 3 and 7 times larger. The ratio of the simulated hit to false alarm runoff is similar in magnitude for the average value (3.5) whereas the median ratio for Lafayette County, 18, is much larger than that seen for the previous EOF site (2.7). In terms of duration separation, the ratios between observed hits and misses, and simulated hits and false alarms, are both smaller for Lafayette County. Both datasets show averages and median duration ratios between 1 and 2.

3.9.3 Manitowoc County Edge-of-Field Runoff Event Comparison

The third EOF site (17.6 acres) located in Manitowoc County is found in the very southern edge of an irregularly shaped NWS basin called WI06C (245,870 acres) where it represents 0.007% of the total area. This NWS basin incorporates a lot of land along the shore of Lake Michigan and drains numerous small streams that empty into the lake. There is no operational river forecasting done in this area therefore this basin was created in the NCRFC system to help generate Flash Flood Guidance (FFG) and given a generic name representing the sixth county basin in Wisconsin.

The analysis for this site produced 62 observed runoff events where 53 where classified as hits and only 9 were missed. That produced an exceptional 85% hit rate and only a 15% miss rate. For the 88 simulated events produced, 46 were hits and 42 were considered false alarms. That breakdown is a 52% hit rate and a 48% false alarm rate. The summary metrics (Table 11) included a Bias score of 1.53 and a CSI of 0.51, both better than the previous two EOF sites. Unsurprisingly, the POD was 0.85 and the FAR was 0.48. Examining the hit-to-miss and hit-to-false alarm ratios provided more good news. For the observed dataset the average hit/miss value was extremely high at 39 where the median was still a very good 14. For the simulated the average value ratio was near 7 whereas the median was around 14. The durations for hit events were generally twice as long as those for misses (observed dataset) and false alarms (simulated dataset).

Table 11. Comparison of runoff events between Manitowoc County EOF and NWS WI06C.

Comparison of Runoff Events between Manitowoc County EOF and WI06C

	Events	Hit	Miss	FA	% Hit	% Miss	% FA
OBS	62	53	9	--	85%	15%	--
SIM	88	46	--	42	52%	--	48%

Observed Event Runoff (in)		**Simulated Event Runoff (mm)**		
Hit Average	0.2537	Hit Average	0.8486	**Bias**
Hit Median	0.0493	Hit Median	0.5384	1.53
Miss Average	0.0065	FA Average	0.1112	
Miss Median	0.0033	FA Median	0.0379	**POD**
Avg Hit/Miss	39.03	Avg Hit/FA	7.63	0.85
Med Hit/Miss	14.94	Med Hit/FA	14.21	

Observed Event Duration (h:m)		**Simulated Event Duration (h:m)**		
Hit Average	22:40	Hit Average	43:18	**FAR**
Hit Median	15:55	Hit Median	27:00	0.48
Miss Average	9:35	FA Average	19:00	
Miss Median	7:30	FA Median	18:00	**CSI**
Avg Hit/Miss	2.37	Avg Hit/FA	2.28	0.51
Med Hit/Miss	2.12	Med Hit/FA	1.5	

3.9.4 Kewaunee County Edge-of-Field Runoff Event Comparison

Representing only 0.02% of the total area of the KEWW3 basin modeling the Kewaunee River near Kewaunee, Wisconsin (81,280 acres), the final EOF site (13.2 acres) in Kewaunee County is located near the eastern side of watershed as water drains towards Lake Michigan. Similar to the nearby Manitowoc County EOF site, the comparison here indicated great results reflected by 86% of the 95 total observed events classified as hits while the miss rate was only 14%. False alarms were numerous here with a FAR 0.62 producing 88 false alarm events out of the total 142 simulated events. The corresponding simulated hit rate was 38%. The Bias score was 1.79 and the CSI was 0.45 (Table 12).

As with the other three EOF sites examined, the Kewaunee County site also has a stratification between the hit events and the miss events. Here the ratio indicates the hit events are 2 to 3 times larger than the misses depending on using the average or median values. For the simulated dataset the hit events were found to be between 5 and 6 times larger than the false alarm values. The durations of hits to misses were found to be between 1 and 1.5 times longer whereas for the simulated events to false alarms were also found between 1 and 2 times longer.

Table 12. Comparison of runoff events between Kewaunee County EOF and NWS KEWW3.

Comparison of Runoff Events between Kewaunee County EOF and KEWW3

	Events	**Hit**	**Miss**	**FA**	**% Hit**	**% Miss**	**% FA**
OBS	95	82	13	--	86%	14%	--
SIM	142	54	--	88	38%	--	62%

Observed Event Runoff (in)		**Simulated Event Runoff (mm)**		
Hit Average	0.1996	Hit Average	5.2186	**Bias**
Hit Median	0.0997	Hit Median	4.4676	1.79
Miss Average	0.1016	FA Average	1.0197	
Miss Median	0.0337	FA Median	0.6904	**POD**
Avg Hit/Miss	1.96	Avg Hit/FA	5.12	0.86
Med Hit/Miss	2.96	Med Hit/FA	6.47	

Observed Event Duration (h:m)		**Simulated Event Duration (h:m)**		
Hit Average	14:43	Hit Average	41:26	**FAR**
Hit Median	9:08	Hit Median	24:00	0.62
Miss Average	10:04	FA Average	22:25	
Miss Median	8:39	FA Median	18:00	**CSI**
Avg Hit/Miss	1.46	Avg Hit/FA	1.85	0.45
Med Hit/Miss	1.06	Med Hit/FA	1.33	

3.9.5 Summary of Field Scale Comparison

Considering the challenge of validating the RRAF concept based on one farm field, representing on average 0.01% of a modeled watershed, the validation results for the EOF sites are very encouraging. To simplify the summarization of the edge-of-field comparisons, a median of medians approach was adopted and will be the source of the summary statistics in the rest of this section. Essentially the median of the four EOF median values, it is assumed this single value descriptor will more closely resemble the population distribution than using mean values which can be shifted due to very large or very small values in a small dataset.

The observed hit and miss rates for the EOF sites were found to be 80% and 20% respectively. These findings provide confidence in the RRAF approach is applicable to the much smaller field scale. The simulated hit and false alarm rates were found to be 29% and 71%, respectively. The Bias score for the EOF comparison was found to be 2.31 which is much higher than the optimal score of 1.0. In addition, the CSI was found to be 0.35. The exceedance distributions of the simulated hits and false alarms for the four NWS basins used in the EOF comparison are shown in Figures 18 – 21.

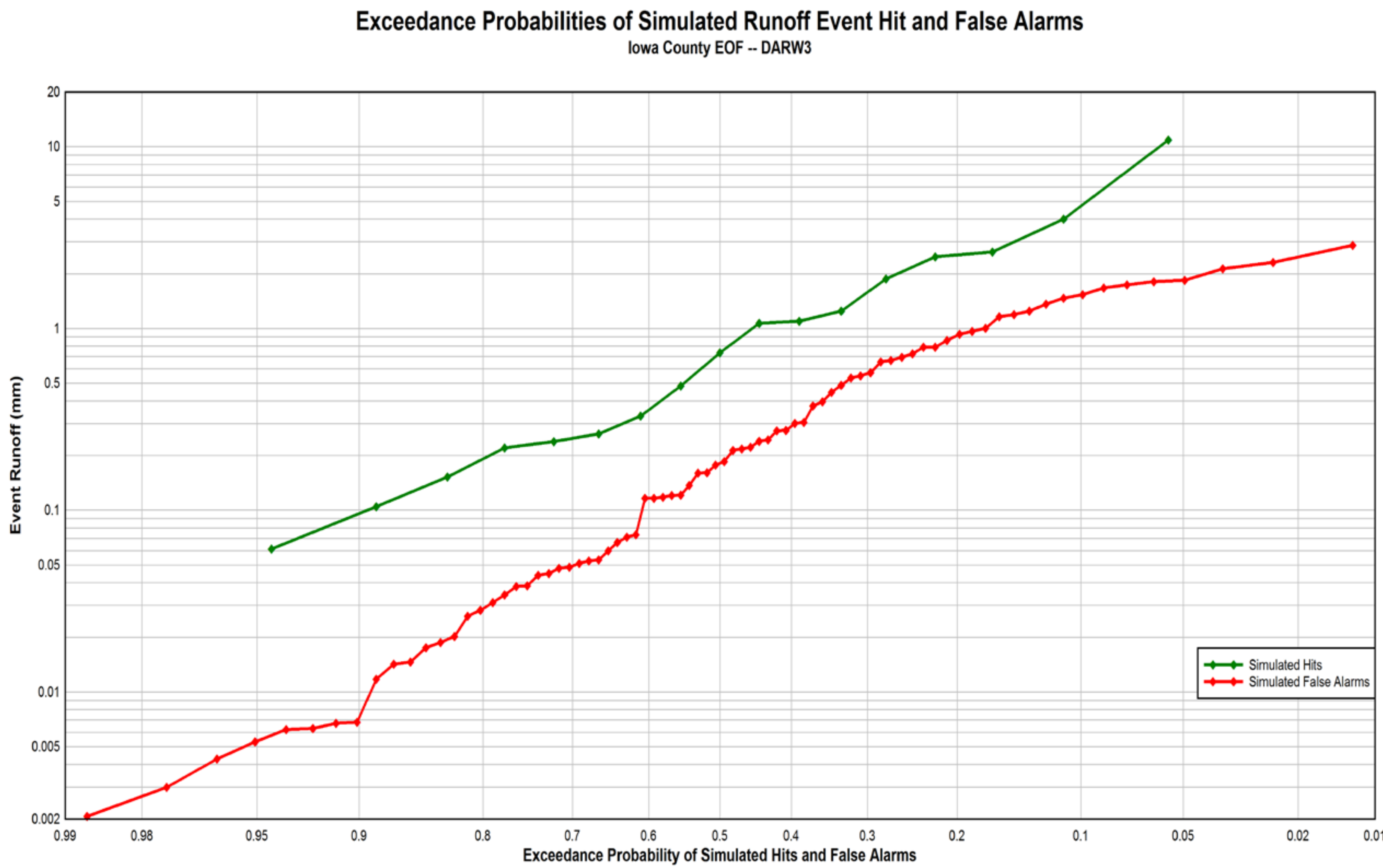


Figure 18. Exceedance probabilities of simulated hits and false alarms for the DARW3 basin.

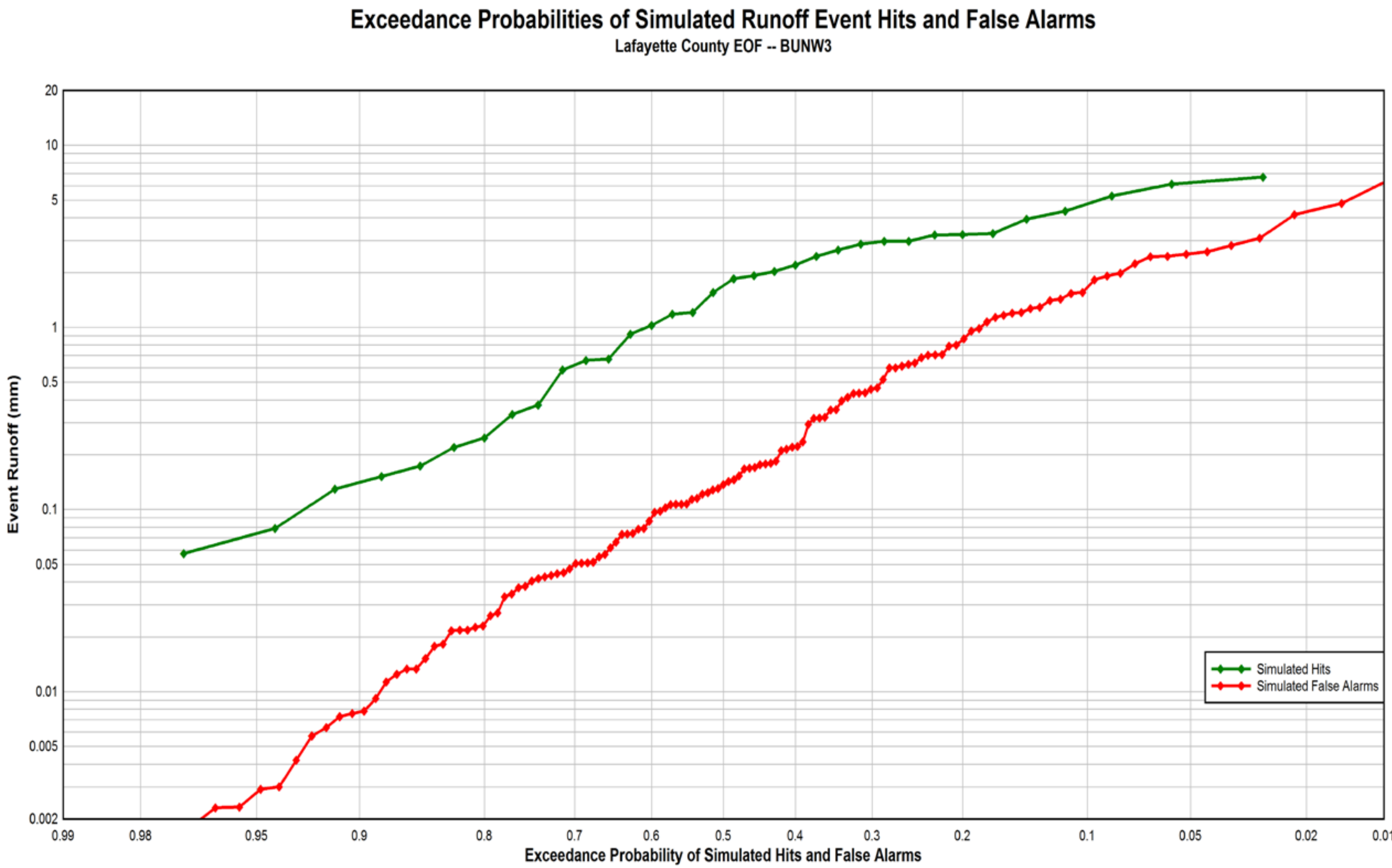


Figure 19. Exceedance probabilities of simulated hits and false alarms for the BUNW3 basin.

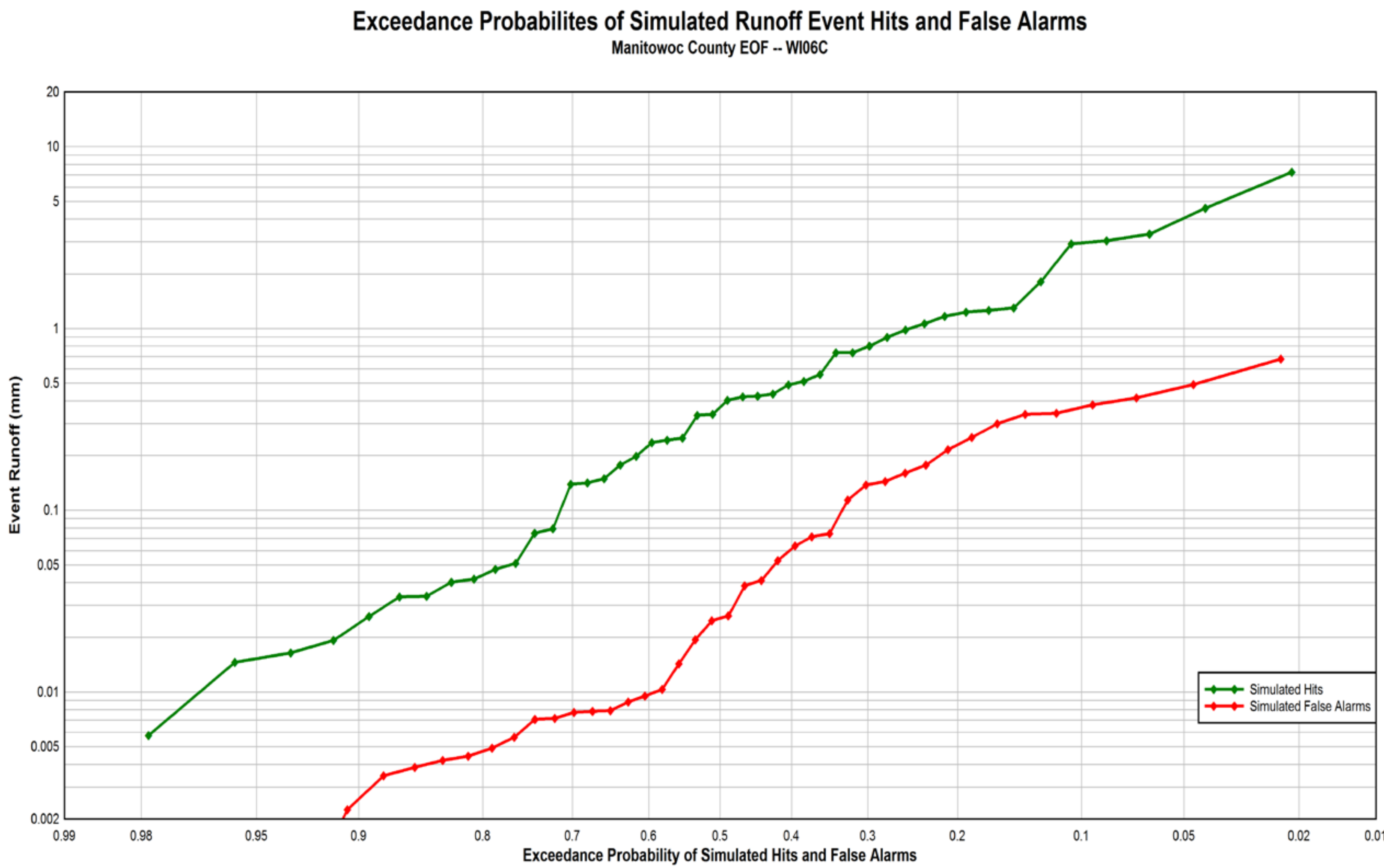


Figure 20. Exceedance probabilities of simulated hit and false alarms for the WI06C basin.

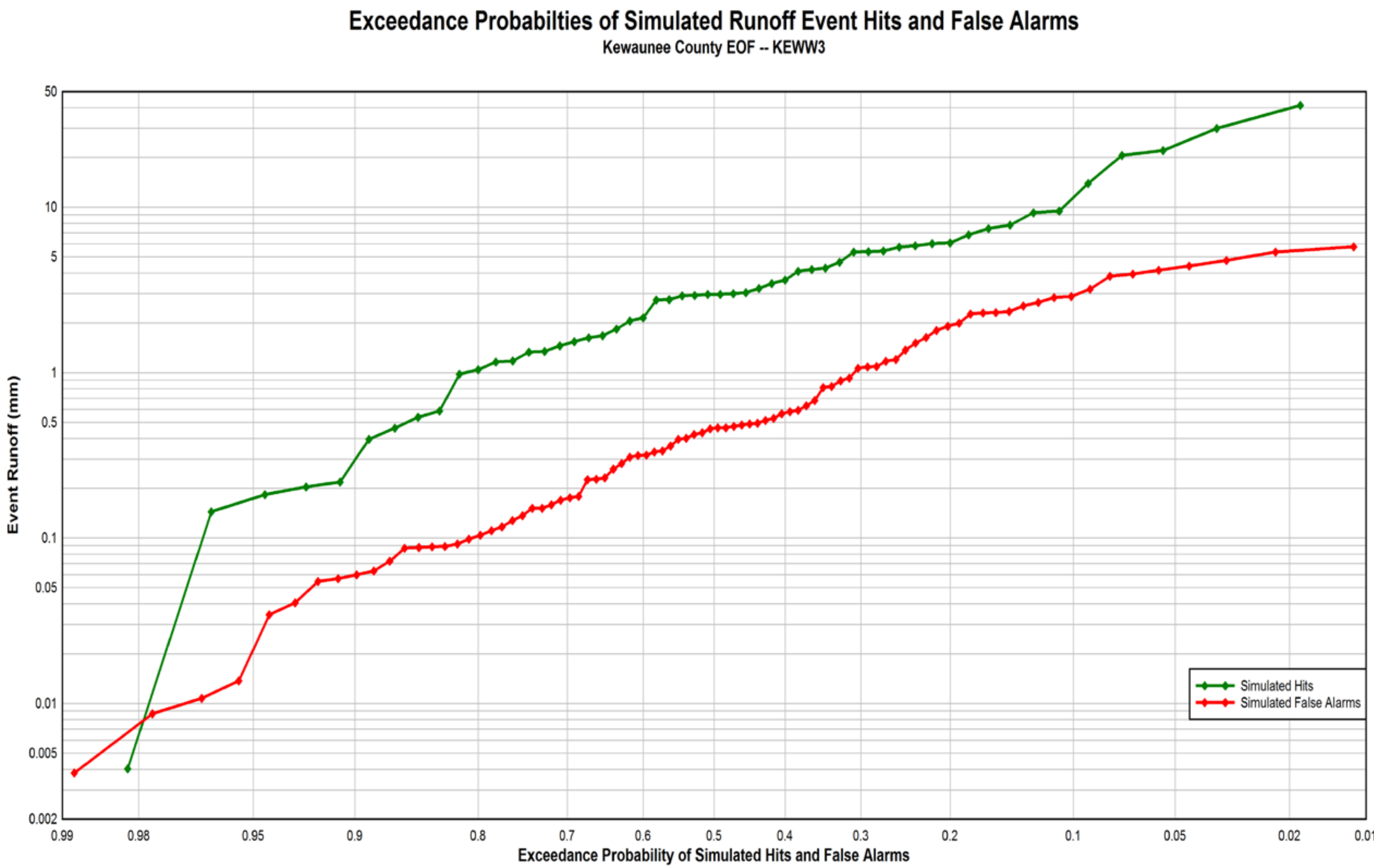


Figure 21. Exceedance probabilities of simulated hits and false alarms for the KEWW3 basin.

As seen in each of the EOF comparisons, examining the differences in both the runoff depth and event duration between the observed hit and miss medians and the simulated hit and false alarm medians suggested more confidence in the RRAF model. The stratification found in the observed dataset further promotes the merit of the RRAF model whereas the simulated dataset suggests further refinement is possible. The median hit-miss ratio for observed data is nearly equal to 5. This indicates the median magnitude of observed hits at the EOF sites is five times larger than the median magnitude of the observed misses. Combining the fact that (1) only 20% of the events were missed, (2) the missed events were much smaller events, and as pointed out in Chapter 2, (3) the smaller events have less risk of carrying contaminants off of fields, it seems reasonable to categorize the observed EOF validation successful.

The difference between the simulated median hit and median false alarm magnitudes was found to be just over 10 for the EOF sites. While the high FAR reported earlier is reason for concern, the hit-miss ratio of 10 suggests refinement is very possible via a threshold to disregard the smaller events which are predominately categorized as false alarms. The ratios of hit-to-miss and hit-to-false alarms in terms of event duration were both found to be near 1.5.

3.10 Small Watershed Runoff Event Comparison

The number of observed runoff events in the small USGS watersheds ranged from 75 in the Black Earth Creek watershed up to 273 in Jackson Creek (Table 13). Although less than the EOF comparison (80% hit rate), the median percent of observed runoff events classified by hits for the seven USGS watersheds was still a respectable 62%. Bower Creek was the worst performer with a hit rate of only 48%, whereas Black Earth Creek and Little Plover River led the comparison with 75% and 77% respectively. Also noteworthy is that the miss rate for the observed events for the small watersheds did show an increase when compared to the EOF sites, 38% from 20%.

The median percent of simulated hits for the watersheds was 55% with the false alarm rate 45%. This is a decent improvement over the 29% EOF median simulated hit rate as well as a decrease in the FAR from 0.71 to 0.45. The best performers in terms of simulated hit rate were Otter Creek, Jackson Creek, and Little Wolf River with values of 79%, 83%, and 84% respectively. Alternatively, Black Earth Creek (23%), Bower Creek (33%), and Eagle Creek (35%) produced poor simulated hit rates and thus the highest false alarm rates. Interestingly Black Earth and Eagle Creeks did well in the observed runoff comparison, whereas Bower Creek was a poor performer in both. The small watersheds' median Bias score and CSI were found to be 2.06 and 0.34 which is consistent with the EOF results.

Table 13. Summary of Small USGS Watershed Validation. Note event runoff values for observed dataset are in inches and simulated dataset is in mm.

	Otter Creek	Eagle Creek	Black Earth Creek	Bower Creek	Jackson Creek	L Plover River	L Wolf River
NWS Basin	SEBW3	WIDM5	BLEW3	GREW3	CLIW3	WIRW3	**ROYW3**
Obs Events	228	191	75	110	273	229	84
Obs Hits	141	132	56	53	160	176	49
Obs Miss	87	59	19	57	113	53	35
% Events (Hit / Miss)	62 / 38	69 / 31	75 / 25	48 / 52	59 / 41	77 / 23	58 / 42
Sim Events	469	741	550	441	567	662	170
Sim Hits	372	262	128	144	470	366	142
Sim FA	97	479	422	297	97	296	28
% Events (Hit / FA)	79 / 21	35 / 65	23 / 77	33 / 67	83 / 17	55 / 45	84 / 16
POD / FAR	0.62 / 0.21	0.69 / 0.65	0.75 / 0.77	0.48 / 0.67	0.59 / 0.17	0.77 / 0.45	0.58 / 0.16
Bias / CSI	1.04 / 0.43	3.20 / 0.20	6.37 / 0.11	3.18 / 0.13	0.94 / 0.43	2.06 / 0.34	0.92 / 0.44
Event Runoff							
Obs Hit Avg	0.3966	0.1152	0.1311	0.7424	0.5454	0.0517	0.2969
Obs Miss Avg	0.0186	0.0179	0.0172	0.1113	0.0370	0.0110	0.0238
Hit/Miss Ratio (Avg)	21.32	6.44	7.62	6.67	14.74	4.70	12.47
Obs Hit Med	0.2056	0.0690	0.0709	0.4076	0.4678	0.0303	0.1380
Obs Miss Med	0.0094	0.0095	0.0107	0.0017	0.0120	0.0080	0.0132
Hit/Miss Ratio (Med)	21.87	7.26	6.63	239.76	38.98	3.79	10.45
Sim Hit Avg	1.2290	1.3856	0.8529	0.6890	1.0117	2.9747	1.5596
Sim FA Avg	0.2350	0.6384	0.5737	0.3485	0.4707	0.4432	0.8835
Hit/FA Ratio (Avg)	5.23	2.17	1.49	1.98	2.15	6.71	1.77
Sim Hit Med	0.7627	0.9822	0.5136	0.1288	0.4724	1.5947	1.0864
Sim FA Med	0.0872	0.1583	0.2174	0.0871	0.0683	0.2579	0.4910
Hit/FA Ratio (Med)	8.75	6.20	2.36	1.48	6.92	6.18	2.21

As with the EOF comparison, the small watershed comparison found beneficial differences between the runoff events classified as hits and either misses or false alarms. The median ratio for the observed hits-to-misses ended up being 10.45 for the small watersheds, up from near 5 for the EOF sites. The median ratio for simulated hits-to-false alarms decreased some from the EOF sites (10.34), but was still a healthy 6.18. The median ratios for the event duration indicated that the observed hits are generally twice as long as the observed misses. However, for the simulated hits and false alarms, the ratio is 1 as the median values of both categories for all the watersheds were 18 hours.

Increasing the amount of NWS basin area represented by the observed watershed from 0.01% to 5.4% with the small USGS watersheds still leads to similar conclusions as derived from the EOF study. The RRAF model and approach seem to be doing a decent job of capturing observed events at both scales, with oddly the smaller EOF sites the better of the two (80% to 62%). Increasing the size of the observed watersheds did decrease the false alarm rate by 37% (0.45 from the EOF value of 0.71). Although the observed miss rate did increase 18% in the watershed comparison (38% from 20%), the value of the median missed event being an order of magnitude smaller than the median hit value helps put that discovery in perspective.

To gain the overall evaluation of the dual-scale RRAF validation, the average of the final median summary statistics was calculated. Note that the ratio values are not calculated from these newly derived average values (Table 14), but are directly the average value of the two observed dataset medians. The exceedance probabilties of the simulated hits (green line) and false alarms (red line) for the seven NWS basins used in the small watershed validation are shown in Figures 22 through 28.

Table 14. The median summary statistics for the two spatial scales were averaged to provide an overall perspective on the RRAF Validation.

Combined Summary Statistics of EOF and USGS Validation			
Obs Hit Runoff (in)	0.11	**Bias**	2.19
Obs Miss Runoff (in)	0.01	**CSI**	0.34
Hit / Miss Ratio	7.71	**POD**	0.71
		FAR	0.58
Sim Hit Runoff (mm)	1.19		
Sim FA Runoff (mm)	0.18	**% Obs Events Hit**	71%
Hit / FA Ratio	8.26	**% Obs Events Missed**	29%
		% Sim Events Hit	42%
		% Sim Events FA	58%

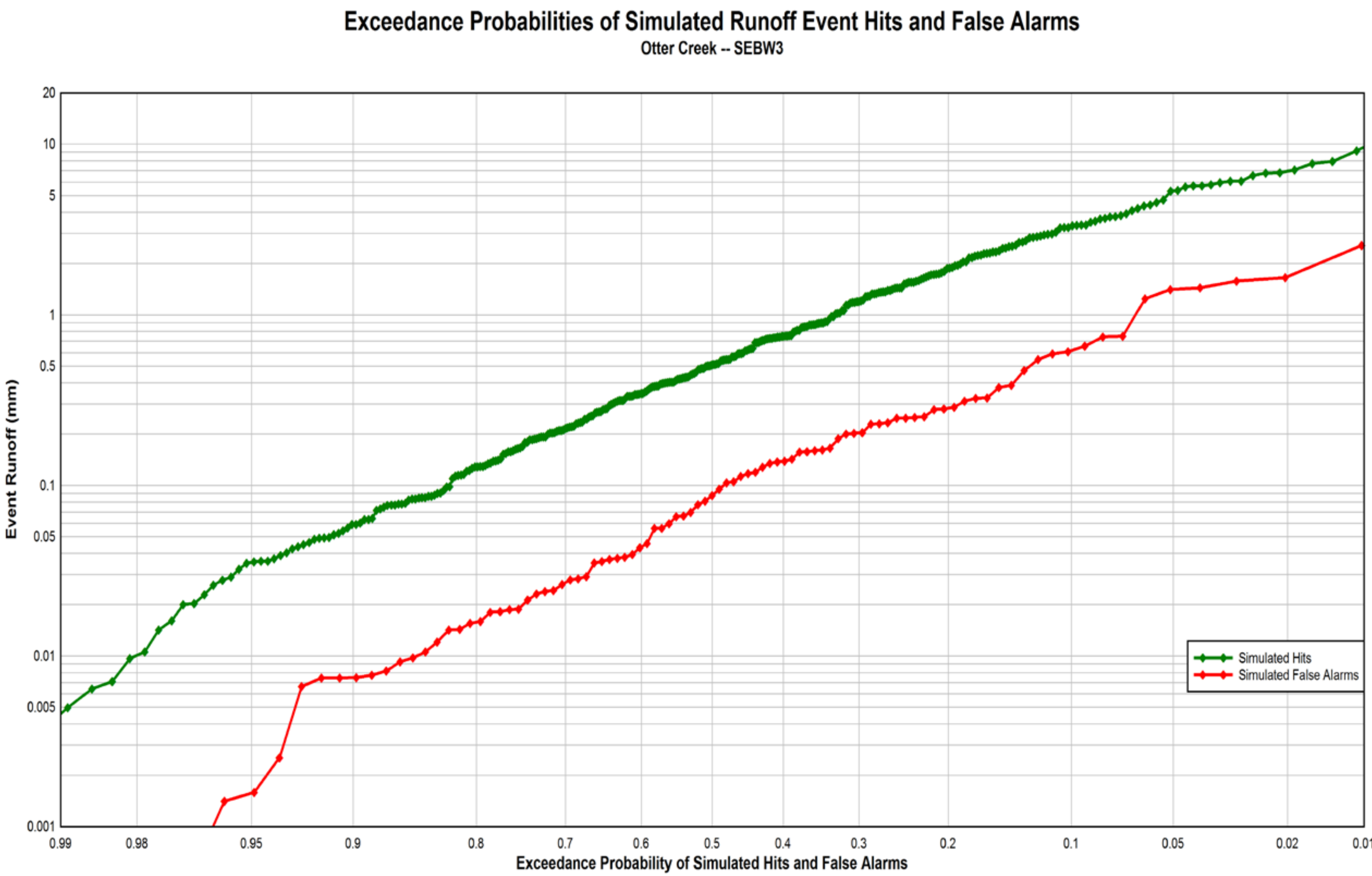


Figure 22. Exceedance probabilities of simulated hits and false alarms for the SEBW3 basin.

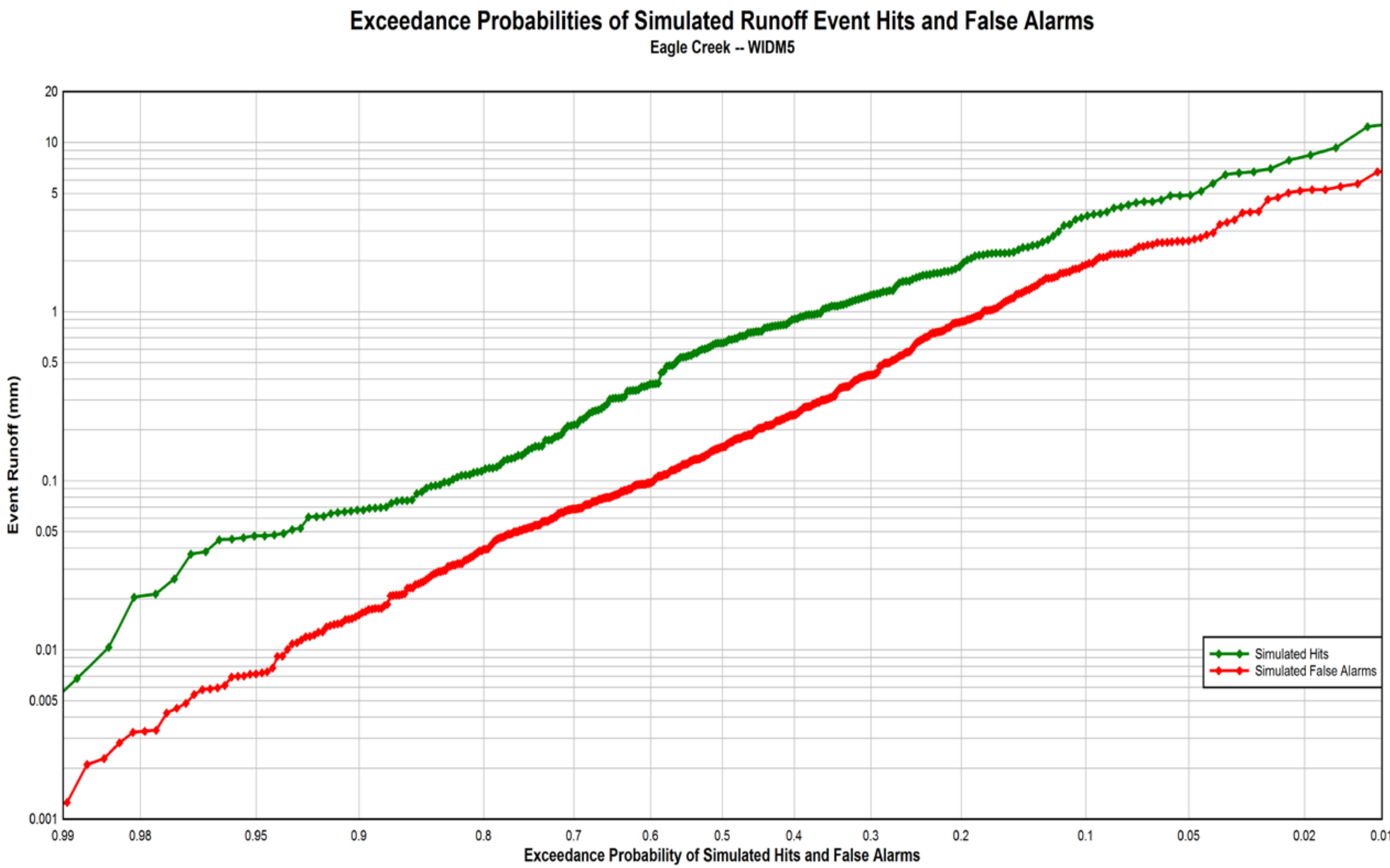


Figure 23. Exceedance probabilities of simulated hits and false alarms for the WIDM5 basin.

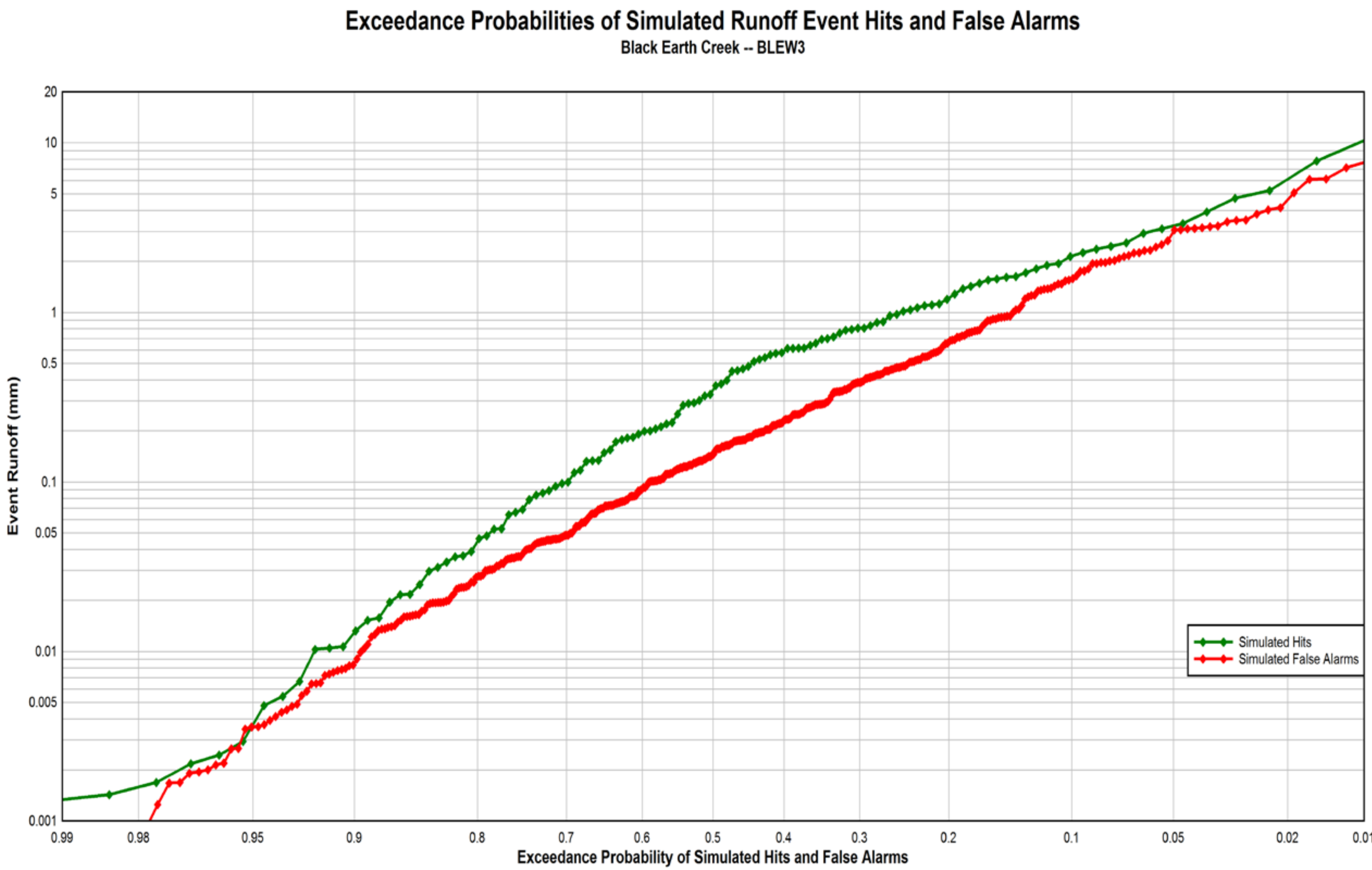


Figure 24. Exceedance probabilities of simulated hits and false alarms for the BLEW3 basin.

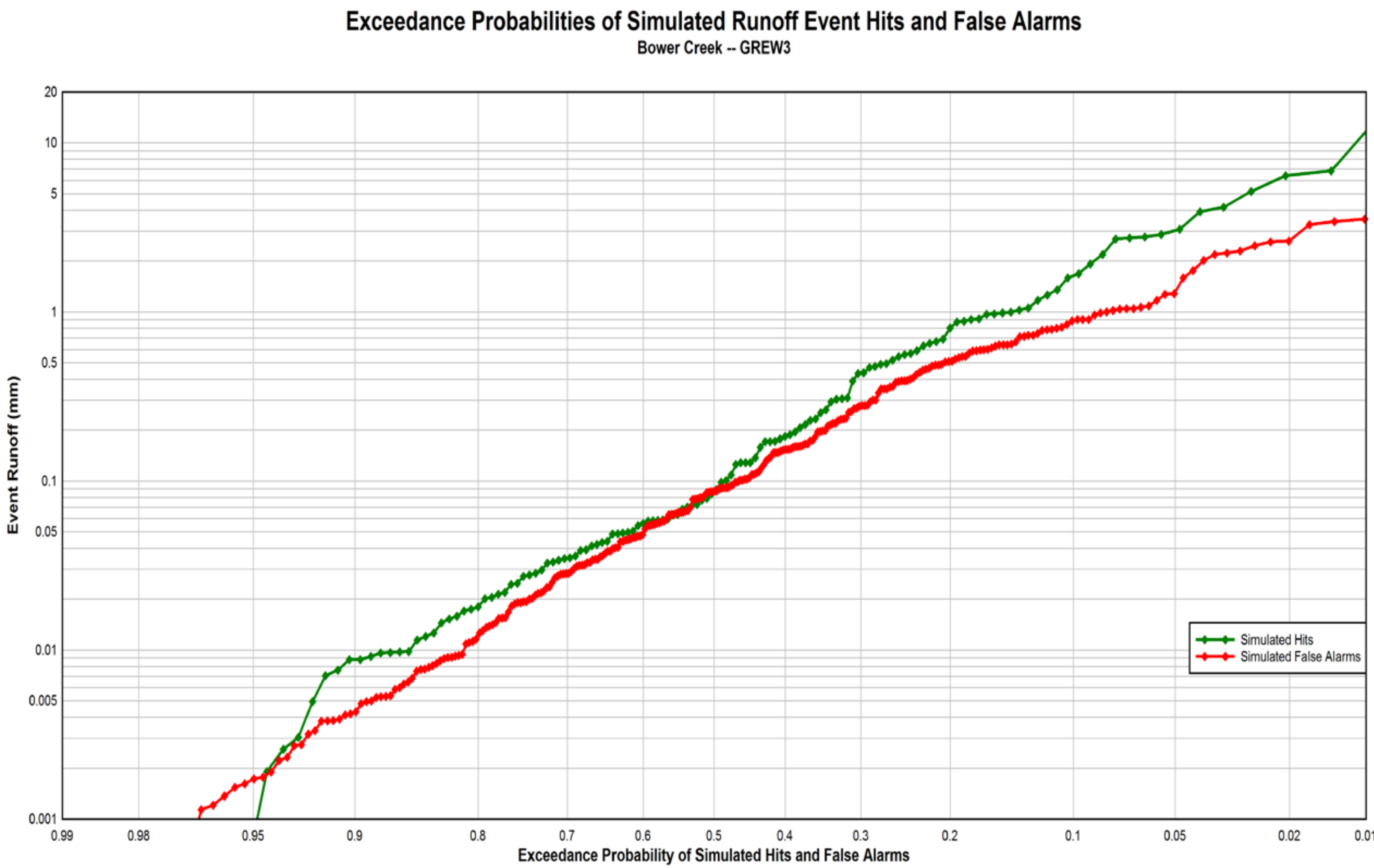


Figure 25. Exceedance probabilities of simulated hits and false alarms for the GREW3 basin.

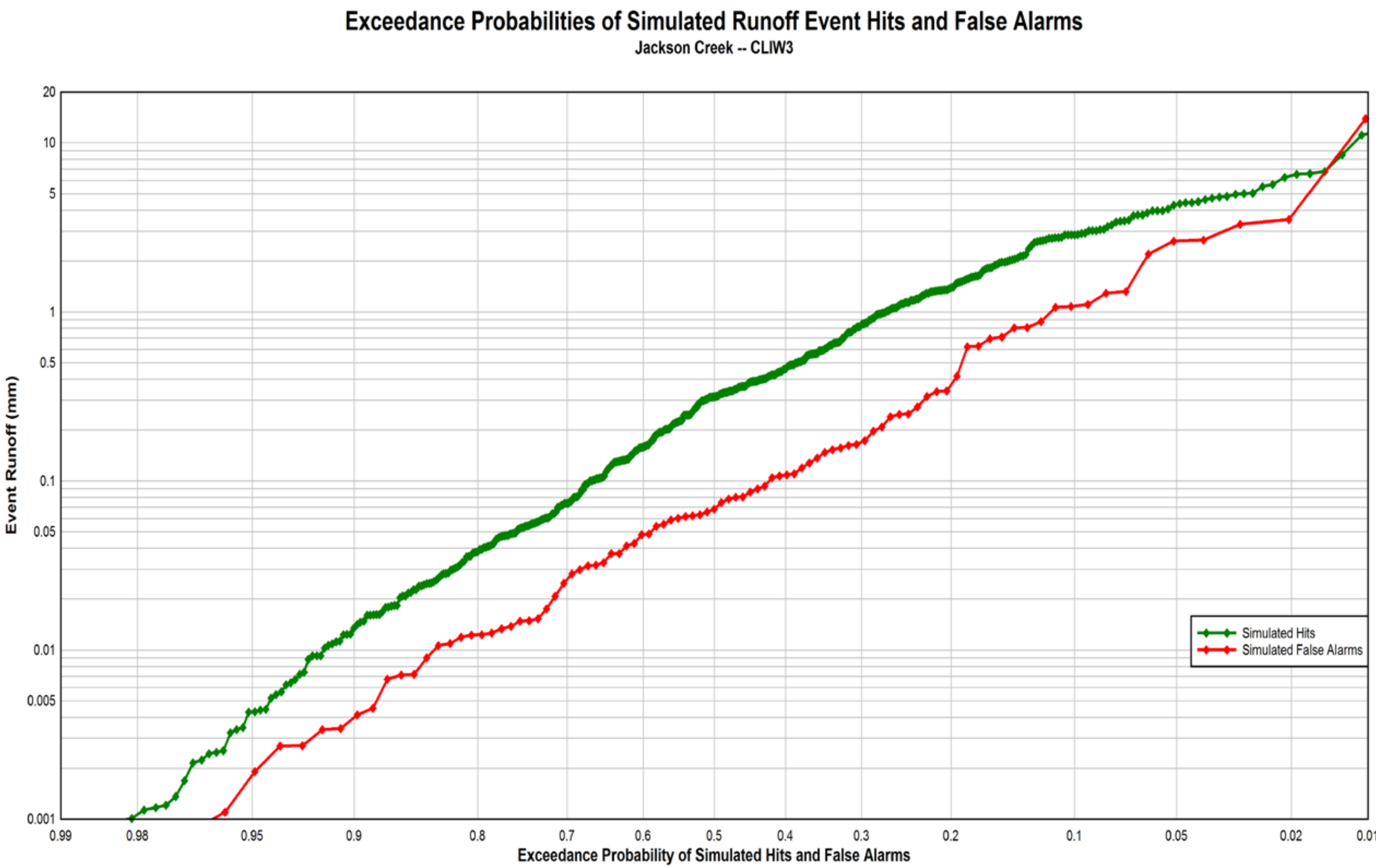


Figure 26. Exceedance probabilities of simulated hits and false alarms for the CLIW3 basin.

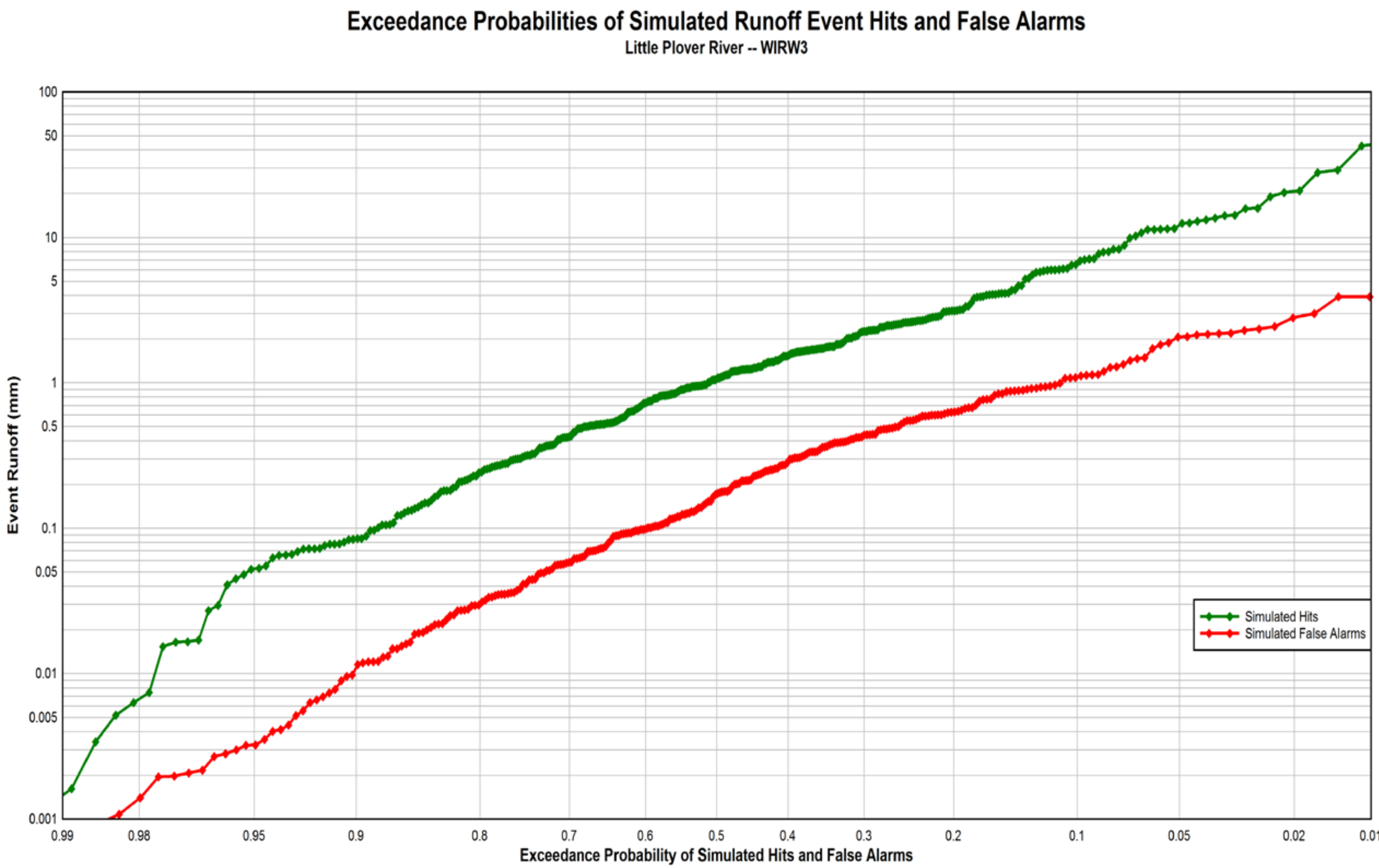


Figure 27. Exceedance probabilities of simulated hits and false alarms for the WIRW3 basin.

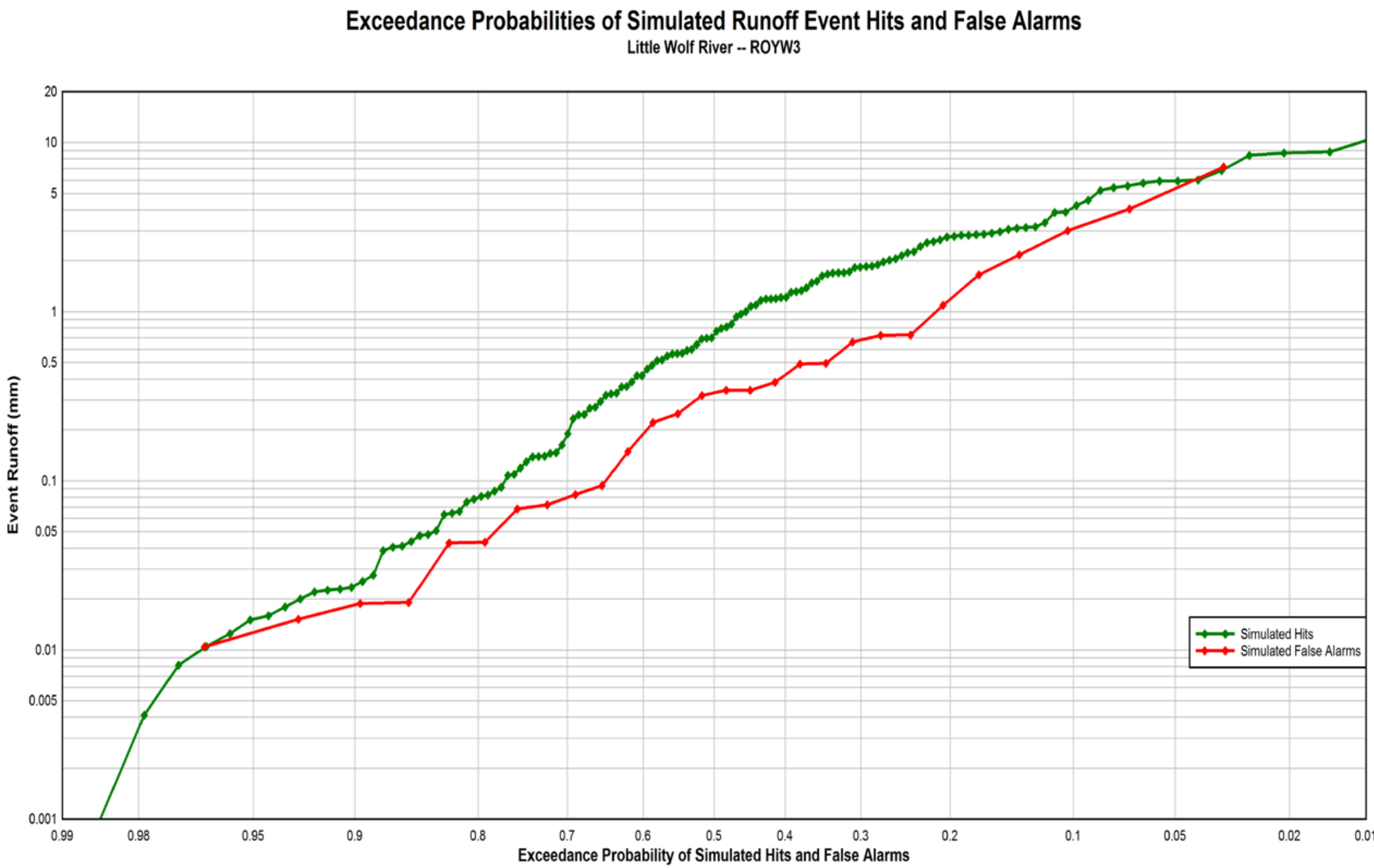


Figure 28. Exceedance probabilities of simulated hits and false alarms for the ROYW3 basin.

3.11 Introduction of Basin Thresholds

As indicated in the previous sections, the RRAF model concept has proven to be a reasonable tool at depicting runoff risk on agriculture fields despite challenging scale differences. However, the validation indicated a weakness of the RRAF appears to be the propensity to generate false alarm runoff events. In any forecasting arena, having a high rate of false alarms is considered very undesirable. Warning too often without the actual hazard occurring promotes "warning fatigue" in the user community. Although the users understand the forecast is a prediction and won't be 100% accurate, consistent false alarms will ultimately lead to degradation in product credibility. If the producers and spreaders refer to the RRAF and it is consistently producing false alarms, they will quickly begin to question and ignore the guidance. The danger with that scenario involves the fact that the RRAF does much better with the major runoff events for a basin. Therefore, if the model produces too many false alarms during smaller runoff risk episodes, the producers might not follow the RRAF during larger and more important events when the guidance is likely to be correct.

An alternate, more optimistic perspective regarding the false alarm issue was presented earlier during the justification of expanding the validation to small USGS watersheds and still applies. By comparing the model to observations representing only 0.01% or 5.4% of the modeled watershed, a glaring weakness becomes apparent in regards to false alarm calculations. It could be argued the false alarm rate is suffering in response to the large percentage of watershed not being measured. Perhaps during many of these false alarm runoff events, other fields or maybe even a majority of the NWS basin is observing runoff, and therefore the event should have been a hit. Thus the validation statistics are a very conservative estimate and are not skewed towards the favor of the RRAF.

Alas, the optimistic suggestion that the false alarms are over-represented can never be more than a theory as current observational datasets can't provide a substantially larger areal coverage for even one of the NWS basins. Therefore, relying on the data that is available, and the risks associated with false alarms, it was decided to proceed with a threshold approach which would be applied to the simulated runoff events to categorize the forecasts into different risk classifications.

3.11.1 Generating Basin Thresholds

Developing a method to produce basin thresholds for the 216 NWS basins presented itself as a challenging task. Preferring not to choose an arbitrary approach, a method derived from the eleven validation basins was needed. Further, the approach had to be universal so that it could be applied to the 205 basins where no validation was completed. After some debate, a suggestion from Brian Connelly (Senior Hydrologist at NCRFC) was used to calculate the basin thresholds. The approach centers around identifying the event runoff value associated with the maximum difference in exceedance between the distributions of simulated hits and false alarms for a basin. This approach essentially maximizes the number of simulated false alarms removed while preserving the maximum number of simulated hits generated. For each runoff value in the overlapping range of the hit and false alarm distribution the exceedance value of the false alarm curve was subtracted from the exceedance value of the hit curve. The differences between the exceedance curves as a function of event magnitude are plotted in Figures 29 and 30.

To produce the final threshold value for each of the NWS basins the following steps were followed:

1. The maximum difference in exceedance between simulated hit and false alarms was found

2. The corresponding event runoff (mm) to the exceedance value of the maximum difference was noted

3. The event runoff from step 2 was cross-referenced with the historical simulations to find the historical distribution exceedance value for that runoff value (Figures 8 and 9).

4. The historical exceedance thresholds for the eleven test basins were evaluated for a pattern which could be applied to the historical distributions of the other 205 NWS basins

The results from generating the thresholds for the eleven test basins are shown in Table 15. The fact that both the average and median exceedance thresholds for both spatial scale validations coalesced around a value of 0.40 was surprising. The watershed comparison basins showed more spread (0.14 – 0.56) than the set used in the EOF comparison (0.36 – 0.51). However, the solid alliance of the mean and median values for both scales, individually and combined, made choosing a universal exceedance threshold of 0.40 an easier decision to justify. The last step in the process consisted of running a program that extracted the threshold runoff value from the historical exceedance distributions for all 216 NWS basins.

The distribution of the basin thresholds is shown in Figures 31 and 32. The median and mean thresholds were found to be 0.4573 mm and 0.5906 mm respectively. The range was found to span from 0.0876 mm to 1.9506 mm.

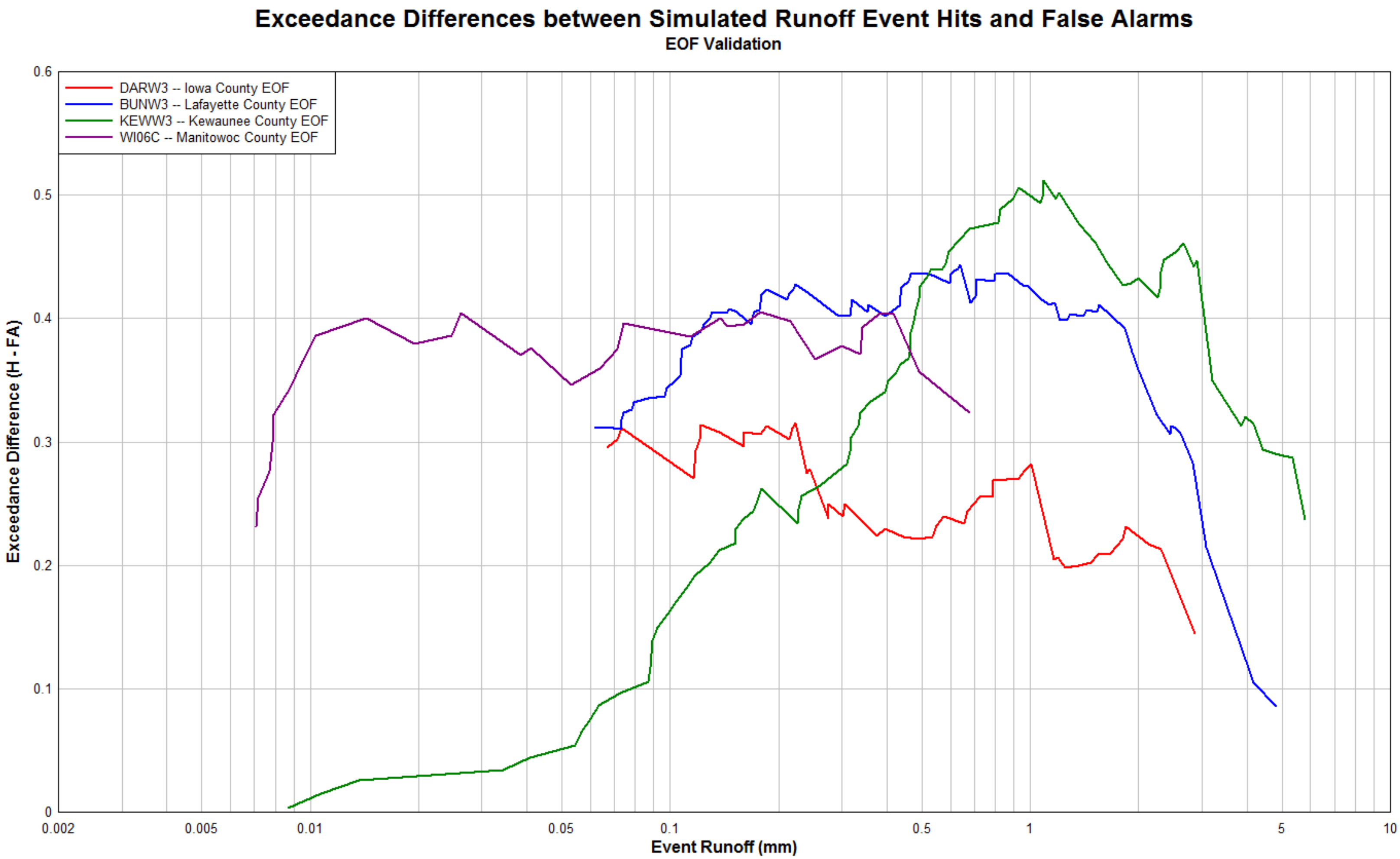


Figure 29. Simulated hit and false alarm exceedance difference for a given event runoff for NWS basins in EOF validation.

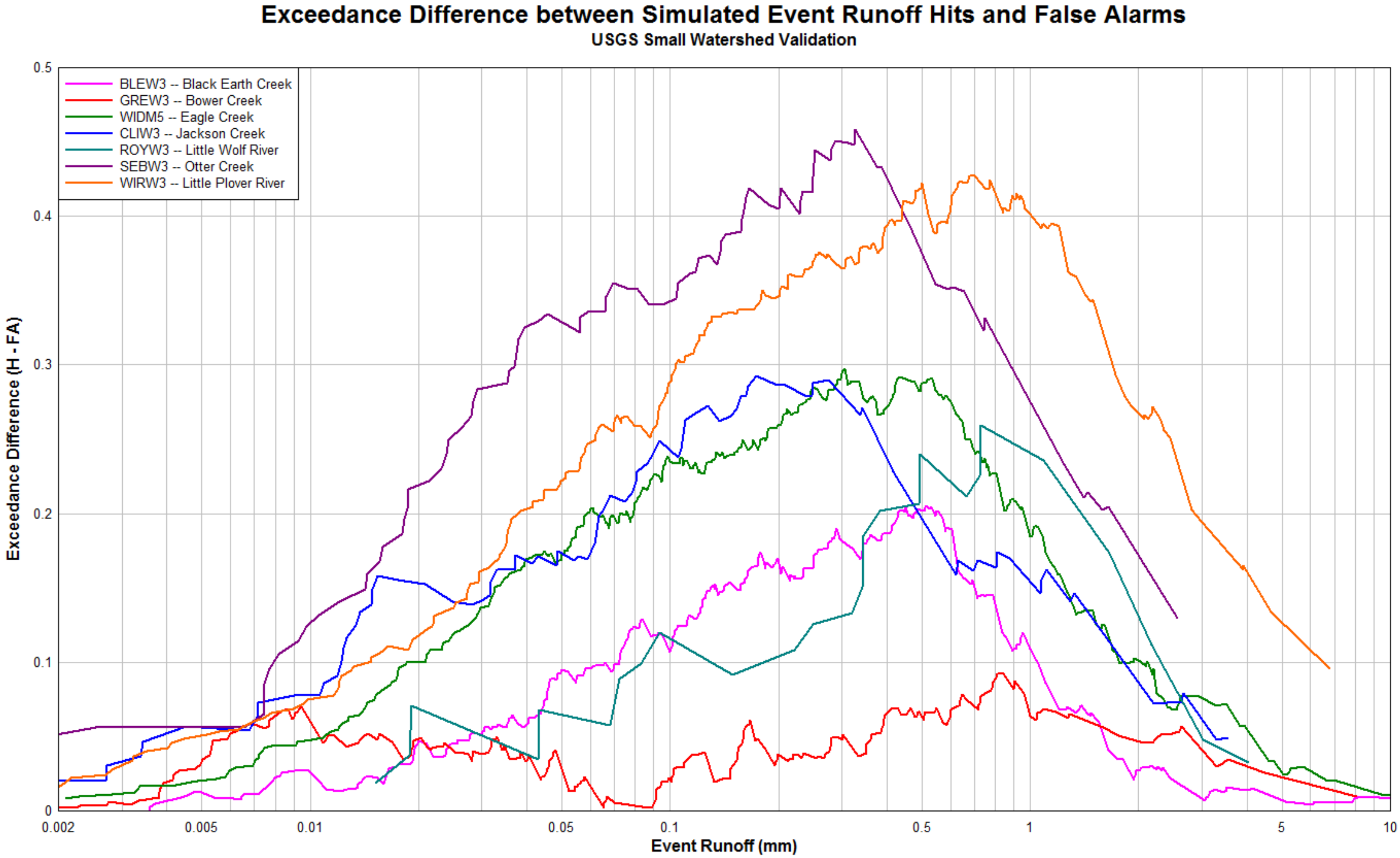


Figure 30. Simulated hit and false alarm exceedance difference for a given event runoff for NWS basins in Small Watershed validation.

Table 15. Calculated data for steps in determining the runoff event threshold based on the eleven validation basins.

	NWS Basin	Exceed Value Max Diff	Corresponding Event Runoff (mm)	Exceed Threshold from Historical Simulation	% of Simulated Historical Events below Threshold
County EOF					
Iowa	DARW3	0.3152	0.2227	0.5135	48.7%
Kewaunee	KEWW3	0.5116	1.0909	0.4078	59.2%
Lafayette	BUNW3	0.4438	0.6384	0.3625	63.8%
Manitowoc	WI06C	0.4051	0.1777	0.3802	62.0%
Average		0.4189	0.5324	**0.4160**	61.0%
Median		0.4245	0.4305	**0.3940**	60.6%
USGS Watersheds					
Black Earth Creek	BLEW3	0.2055	0.5144	0.2657	73.4%
Bower Creek	GREW3	0.0922	0.8430	0.1381	86.2%
Eagle Creek	WIDM5	0.2978	0.3059	0.4006	59.9%
Jackson Creek	CLIW3	0.2925	0.1735	0.5596	44.0%
Little Wolf River	ROYW3	0.2593	0.7289	0.3908	60.9%
Otter Creek	SEBW3	0.4581	0.3252	0.5416	45.8%
Little Plover River	WIRW3	0.4274	0.6770	0.4099	59.0%
Average		0.2904	0.5097	**0.3866**	61.3%
Median		0.2925	0.5144	**0.4006**	59.9%
EOF & Watersheds					
Average		0.3371	0.5180	**0.3973**	60.3%
Median		0.3152	0.5144	**0.4006**	59.9%

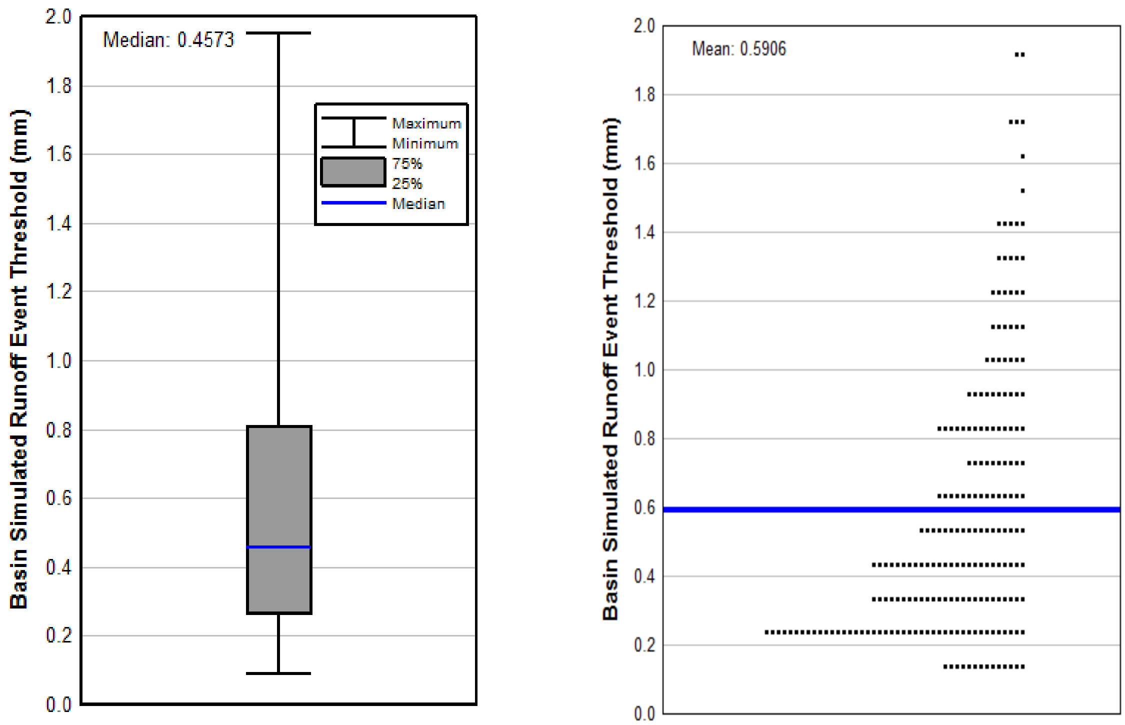


Figure 31. Distribution of the 216 basin runoff event thresholds (mm) generated to reduce false alarms represented in a box plot and histogram.

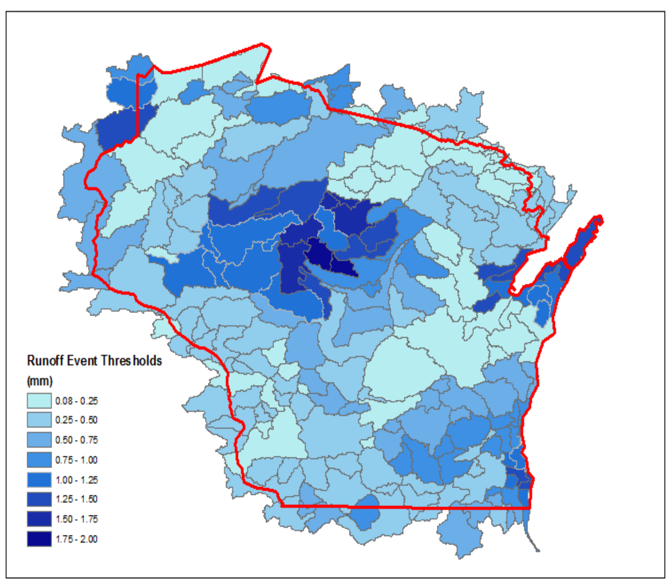


Figure 32. Spatial distribution of basin-specific runoff event thresholds (mm).

3.11.2 Definition of Event Categories

A consequence of applying a threshold to simulated runoff events is the creation of risk categories. By definition every time-step of the forecast will be classified as being one of three possible categories:

- CAT 1 : No event has been simulated
- CAT 2 : An event has been simulated but is less than the basin threshold
- CAT 3 : An event is simulated equal to or greater than the basin threshold

3.12 Field Scale Runoff Event Comparison with Thresholds

To evaluate the effect of the thresholds on the validation performed in earlier sections a second analysis was completed. This examination started by editing the list of historical simulated runoff events for each basin and removing any event below the basin threshold leaving only the newly defined CAT3 events. A comparison with the observed data was then performed and the summary statistics calculated to evaluate the changes due to the thresholds. The differences between the two approaches are summarized in Table 16 where the medians of the EOF validation basin are compared. Values for the non-threshold set were described in section 3.9.5.

Table 16. The effects of the thresholds on the summary statistics of the EOF validation.

Summary of Threshold Application Effect on EOF Validation				
	No Threshold	**Threshold**	**T - NT**	**% Change**
Median Obs Hit	0.0745	0.0961	0.0216	29%
Median Obs Miss	0.0143	0.0195	0.0053	37%
Hit / Miss Ratio	4.96	4.92	-0.04	-0.8%
Median Sim Hit	1.6073	2.3221	0.7148	45%
Median Sim FA	0.2040	1.6539	1.4499	711%
Hit / FA Ratio	10.34	1.41	-8.93	-86%
% Obs Hit	80%	64%	-16%	-21%
% Obs Miss	20%	36%	16%	85%
% Sim Hit	29%	52%	23%	79%
% Sim FA	71%	48%	-23%	-32%
Bias	2.31	1.11	-1.20	-92%
CSI	0.35	0.48	0.13	38%
POD	0.80	0.64	-0.16	-21%
FAR	0.71	0.48	-0.23	-32%

It should be noted that the percent change for the bias metric was computed as a relative change from the perfect score of 1.0 and not strictly from old to new. Overall, introducing the threshold concept continues to produce very acceptable results for the field scale validation. As expected the observed event comparison provided worse results in terms of percent of events categorized as hits and misses as hits dropped 21% and misses increased 85%. However, more importantly, the hit-miss ratio only decreased less than 1% and remains at a value of nearly 5.

Remembering the threshold has no impact on the observed events, finding that the ratio is nearly unchanged highlights that the simulated events below the threshold, even some that were hits, are very small in magnitude and importance.

With regards to the simulated events, the false alarm ratio dropped 32% to a value of 0.48 after the thresholds were imposed confirming the expected goal. A drastic difference in the hit-false alarm ratio also resulted from the threshold application yet was not unexpected as the lower 60% of the simulated runoff events were removed from the distribution pushing the median values of hits and false alarms closer.

The Bias and CSI metrics are useful in evaluating the effect of the thresholds on the RRAF model overall. The trends in these two values due to the thresholds again promote the idea that the benefits outweigh the costs with this approach. Remembering that a perfect score is 1.0, the relative median Bias score has a 92% improvement as it drops from 2.31 to 1.11. The CSI also improves with an increase from 0.35 to 0.48. Summarizing the new approach results for the EOF validation, the false alarms were reduced dramatically, and although some observed events were missed, they were of the a much smaller magnitude than the observed events that were classified as hits.

3.13 Small Watershed Runoff Event Comparison with Thresholds

The same process described in section 3.12 was followed to generate the new dataset needed for evaluating the effects of thresholds on the small watershed validation. The summary of the differences is shown in Table 17.

Table 17. The effects of the thresholds on the summary statistics of the small watershed validation.

Summary of Threshold Application Effect on Small Watershed Validation				
	No Threshold	**Threshold**	**T - NT**	**% Change**
Median Obs Hit	0.1380	0.2179	0.0799	58%
Median Obs Miss	0.0095	0.0120	0.0025	26%
Hit / Miss Ratio	10.45	14.24	3.79	36%
Median Sim Hit	0.7627	2.0031	1.2404	163%
Median Sim FA	0.1583	1.0766	0.9183	580%
Hit / FA Ratio	6.18	1.28	-4.90	-79%
% Obs Hit	62%	41%	-21%	-33%
% Obs Miss	38%	59%	21%	54%
% Sim Hit	55%	81%	26%	46%
% Sim FA	45%	19%	-26%	-57%
Bias	2.06	0.76	-0.82	-77%
CSI	0.34	0.36	0.02	7%
POD	0.62	0.41	-0.20	-33%
FAR	0.45	0.19	-0.26	-57%

The same themes found when applying thresholds to the EOF validation were present in the small watershed analysis. The percent of observed events hit did decrease and the misses did increase. Although the hit-miss ratio for the EOF comparison remained nearly unchanged, the ratio for the small watersheds actually increased as very small magnitude hits were removed from the distribution. As before, this combination indicates that although the percent of observed events hit is lower, the threshold application enables the RRAF to do well with the larger magnitude, more important events.

The FAR score had a dramatic 57% decrease from 0.45 to 0.19 in this analysis. The simulated hits saw a corresponding increase of 46% jumping from only 55% to 81% after thresholds were applied. As with the EOF comparison, it seems the threshold approach is working in reducing false alarms while focusing the emphasis on the more important runoff events. The Bias score supports this with an improvement of 77% while the CSI only indicated a much smaller improvement of only 7%.

3.14 Combined Field Scale and Small Watershed Comparison

Following along the method described in section 3.10 (Table 14) the average of the two spatial scale analyses was calculated to provide an overall review of the RRAF model after thresholds were applied. The summary is shown in Table 18 below.

Table 18. The effects of thresholds shown on combination of EOF and small watershed summary statistics.

Summary of Threshold Application Effect on Average Values of Combined EOF and Small Watershed Validation				
	No Threshold	**Threshold**	**T - NT**	**% Change**
Avg Obs Hit Medians	0.11	0.16	0.05	48%
Avg Obs Miss Medians	0.012	0.016	0.004	33%
Hit / Miss Ratio	7.71	9.58	1.87	24%
Avg Sim Hit Medians	1.19	2.16	0.98	83%
Avg Sim FA Medians	0.18	1.37	1.18	654%
Hit / FA Ratio	8.26	1.35	-6.91	-84%
% Obs Hit	71%	53%	-18%	-26%
% Obs Miss	29%	47%	18%	64%
% Sim Hit	42%	66%	24%	58%
% Sim FA	58%	34%	-24%	-42%
Bias	2.19	0.93	-1.12	-94%
CSI	0.34	0.42	0.08	23%
POD	0.71	0.53	-0.19	-26%
FAR	0.58	0.34	-0.24	-42%

Unsurprisingly, averaging the median summary metrics of the two spatial scales resulted in very similar findings to those presented earlier for the EOF and small watershed scale individually. The percent of observed events was very good at 71% with no threshold, however, dropped to only 53% after the application. Conversely, the missed observed events did increase 64% up to a value of 47% missed with the thresholds. Yet, as seen earlier, the consistent ratio difference between the hit and miss event magnitudes claims a more important role in the evaluation of the threshold effectiveness. In fact, the hit-false alarm ratio actually increases due to the threshold application, again supporting the assumption that although a few more events are missed due to the thresholds, those events remain in the very small end of the spectrum.

The simulated event analysis provides the same story as well. The percent of simulated hits increases and the false alarms decrease. The FAR dropped 24% to 0.34 from 0.58 with no thresholds. Further evidence for the improvement of the RRAF model due to the threshold application is provided by the Bias score. With no threshold, the overall value was 2.19. After thresholds were applied on the basins, the score dropped down to 0.93 on a scale where 1.0 is a perfect simulation. This is a 94% improvement for the RRAF model over the original procedure where many false alarms were dampening the effectiveness. The CSI also showed improvement, jumping 23% to a value of 0.42.

3.15 Summary of Validation Results

The validation of the RRAF model against observed runoff data was conducted at two spatial scales as well as before and after the 0.40 exceedance threshold was introduced. Overall, the RRAF concept appears to have done reasonably well considering the inherent scale limitations. The field scale stretches the validation to the extreme, yet is also the scale where the guidance will be ultimately tested and judged by the end users. It just happened that the EOF analysis is where the highest hit percentage of observed events was found with 80% without thresholds and 64%

with. For the small watershed comparison, the values were slightly lower with 62% and 41% (no threshold and threshold respectively). Combining the two scales provided values of 71% with no thresholds and 53% with.

Chapter 3 proposed that applying the thresholds improved the RRAF product. Relying only on the percent observed events classified as hits would incorporate doubt into that suggestion. However, at both spatial scales and regardless of thresholds, there was a very distinct separation of the median runoff event categorized as a hit and miss. This meant that the events were missed were very small in magnitude, and applying the threshold did not change that fact. As discussed in Chapter 2, usually the largest runoff events are the ones that cause the most degradation to nearby water bodies. Therefore, missing the occasional small event as an expense towards substantial reduction of false alarms is a trade off easily accepted by the RRAF team.

The high false alarm rates initially witnessed in the EOF validation were the primary reason to include a comparison against the small watershed data. It was proposed that including the watersheds gauged by the USGS would provide a more robust validation against the NWS basins as they incorporated more than just a single field as in the EOF comparison yet were still small enough to respond similarly to the field scale. A decrease in the false alarms was indeed found for the small watersheds as the FAR scores were lower than the EOF in all cases. While the EOF values were 0.71 and 0.48, the USGS watersheds were 0.45 and 0.19 (no threshold and threshold). The combined FAR scores were 0.58 before and 0.34 after applying thresholds.

Overall metrics such as the Bias score suggest that the RRAF can be a very competent tool on average. Though the validation is based only on eleven test basins, Chapter 3 showcases an extensive analysis that was completed to evaluate the usefulness of the NWS lumped SAC-SMA model for forecasting field scale runoff risk. A final Bias score of 0.93 was found which implies a capable model has been developed. Chapter 4 will highlight the behavior of the RRAF product over an entire year.

4. RESULTS

It goes without saying that the major focus and concern regarding the RRAF will center on how well it works. Product stability, reliability, and accuracy are attributes both manure producers and state agencies will place significant importance and attention on. Chapter 4 will review the most complete summary of the RRAF performance available at this time. RRAF performance will be evaluated for the 2011 calendar year using two approaches for verification. The first examined the spatial trends and inconsistencies in the summation of forecast guidance over the year (sections 4.1 – 4.3). The second consisted of actual verification against additional edge-of-field runoff datasets similar to what was used in the validation process (section 4.4). Section 4.5 will provide a final summary on the RRAF performance relying on validation and verification results. Section 4.6 will document NWS model changes made to optimize the RRAF guidance while some future goals and proposals for the RRAF are detailed in section 4.7.

4.1 2011 Forecast Guidance Analysis Approach

This analysis will incorporate 365 daily morning forecast runs (T_0 = 12Z) from 2011. For each of the 216 NWS basins evaluated, the forecast data (10-day forecast window) was summed to create basin Analysis Accumulations (AA) of various model variables as well as simulated events. It is important to emphasize that only forecast data are being analyzed in this evaluation. A given calendar day will be simulated multiple times in this approach. Further, future storm systems or important snowmelt events will be tallied several times as it occurs in successive forecast runs. Therefore, it is important to stress that Analysis Accumulations are not synonymous to calendar year annual totals.

The analysis will evaluate RRAF product consistency over the year quantitatively as well as spatially. The analysis will enable several questions to be answered such as "how often will my area be classified as high risk of runoff?" Areas of improvement in the underlying models or RRAF concept will also be investigated during this analysis. For example, outlier basins might require model calibration adjustments or simply a change in the basin threshold.

4.2 Analysis Accumulation of Model Parameters

This section will investigate the analysis accumulation of the model variables used in the RRAF product for 2011. These variables include the five days of forecast precipitation (QPF) used in every run, the Rain+Melt (RAIM) time-series from the Snow17 model, and the SAC-SMA components of Interflow Runoff (INTRO) and Upper Zone Tension Water Deficit (UZTWD).

4.2.1 Forecast Mean Areal Precipitation (FMAP)

This analysis consists of summing the 365 daily run totals of five days of QPF, or the FMAP, for each basin. The literal amount of precipitation is not the focus, but instead the concern is on the spatial distribution and possible quantitative anomalies. The analysis found that days two and three had the highest accumulations of QPF while days four and five the lowest (Table 19 and Figure 33).

Table 19. Analysis Accumulated daily and total FMAP for all NWS basins for 2011.

Analysis Accumulated Daily and Total FMAP for all NWS Basins in Wisconsin				
FMAP (mm)	**Median**	**Maximum**	**Minimum**	**Med % Total QPF**
Day 1	927	1220	792	19.7%
Day 2	1032	1288	875	22.1%
Day 3	1006	1208	839	21.6%
Day 4	863	1059	712	18.2%
Day 5	844	1093	744	18.4%
Total	4650	5857	3972	

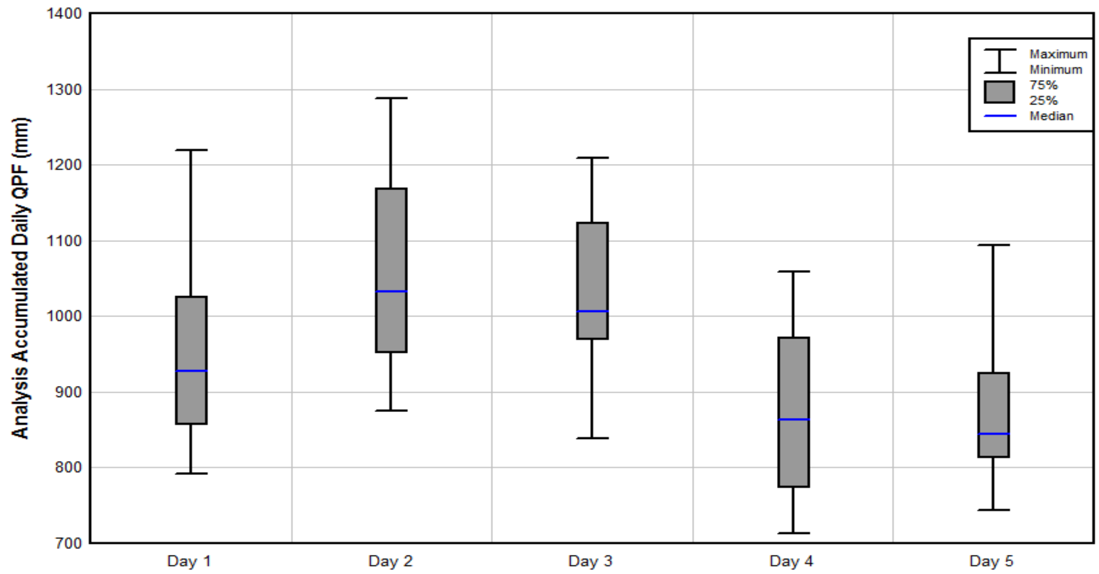


Figure 33. Distribution of Analysis Accumulated daily FMAP totals for all of the NWS basins in Wisconsin.

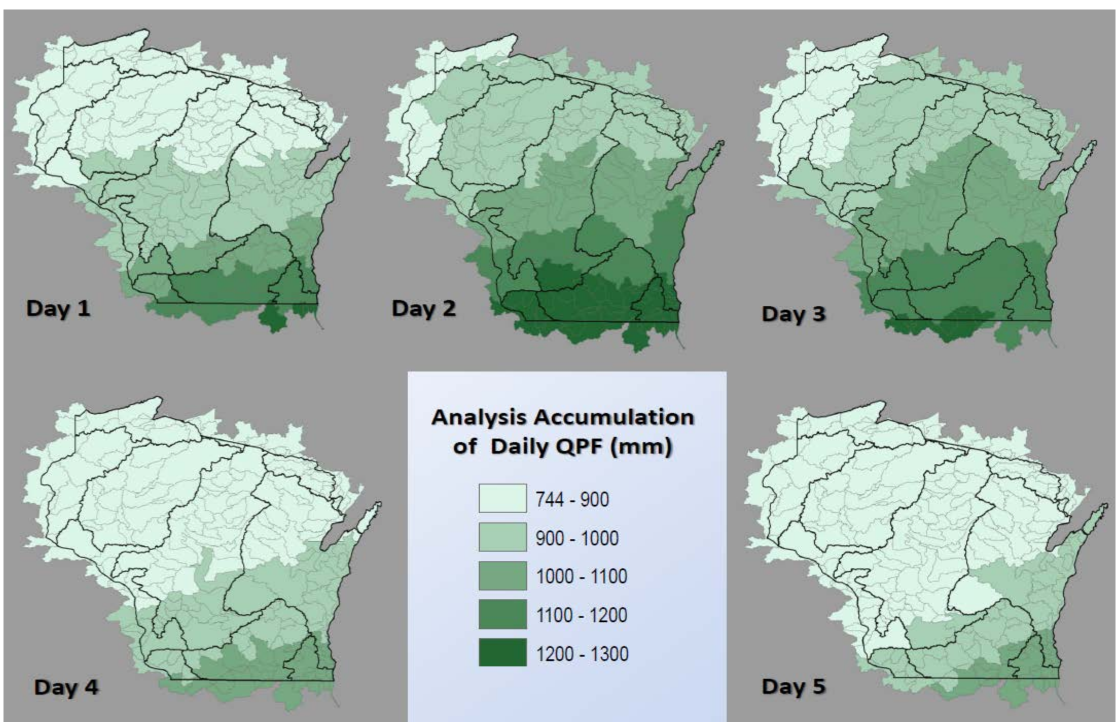


Figure 34. Spatial distribution of Analysis Accumulated daily FMAP (mm) for 2011.

The spatial distribution of precipitation used in the forecasts (Figures 34 and 35) indicates the southern quarter of Wisconsin received the highest totals of FMAP over the year. When examining how often QPF existed in the forecast period it was found that days 4 and 5 recorded the highest percentage for median number of daily runs with FMAP greater than zero. Specifically, for all of the NWS basins, the median percent of daily runs with some QPF was 47%, 49%, and 47% for Days 1 – 3 respectively. Days 4 and 5 had median values of 54 and 55%. This finding corresponds with common knowledge regarding QPF where often days 4 and 5 are produced with very light but widespread forecast precipitation.

A broader look at the existence of daily forecast runs with any amount of forecast QPF is highlighted in Figure 36. The northern edge of Wisconsin had the highest percentage of daily runs with FMAP greater than zero yet is also located in the area of lightest accumulation of FMAP over the year. The median percentage of all the NWS basins in this study indicated that half of the runs contained QPF, while the range of percentages was from 47% to 58%.

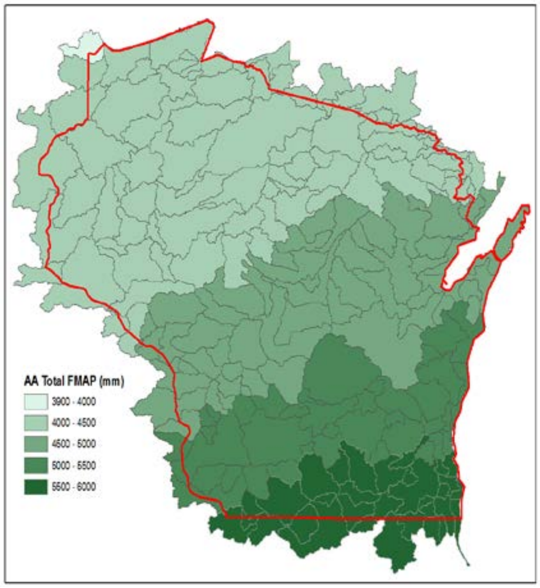


Figure 36. 2011 Analysis Accumulated total FMAP.

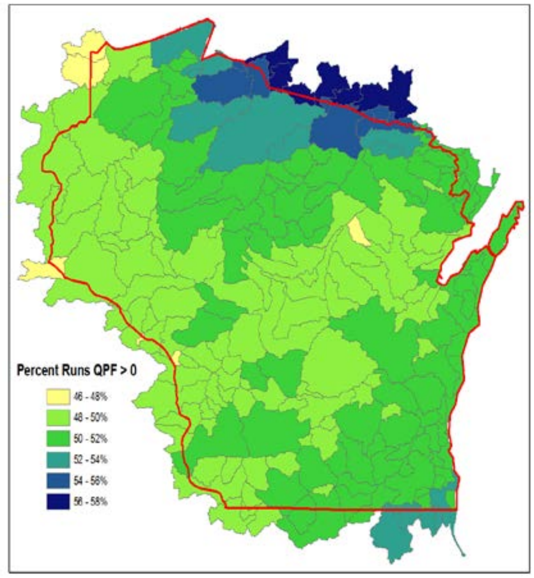


Figure 35. Percent of daily forecast runs in 2011 with forecast precipitation greater than zero.

4.2.2 Snow-17 Rain + Melt (RAIM)

The Rain+Melt (RAIM) time-series created by the Snow-17 model is the next variable evaluated. Analyzing the RAIM component will be very similar to the FMAP with the difference being the incorporation of melt water from the basin snowpack. The accumulation of 14,600 RAIM values for every basin is shown in Figures 37 and 38 (365 daily runs x 40 time-steps). The median AA RAIM total was 5,266 mm with a range between 4,089 and 5,997 mm for the 216 basins. Comparing the AA sums of RAIM and FMAP for all 216 basins found that there was 8% more RAIM produced. However, for twenty five basins the FMAP totals were actually higher than the RAIM totals.

These findings are not that surprising for a couple reasons. These values represent only forecast data which can interject changes in precipitation typing (snow or rain), snowpack simulations, and melt from run to run. In addition, on January 1st, the starting snowpack could be greater than zero for many basins resulting in moisture in the analysis time period not accounted for by the FMAP. Successive daily runs where the model is melting that extra snow would continue to tally in the AA total RAIM yet not increase the FMAP values. Another reason for the difference is due to the RRAF being derived from an operational model. NCRFC forecasters routinely evaluate a variety of data sources and can modify the modeled snowpack moisture by increasing or decreasing it to align with field observations.

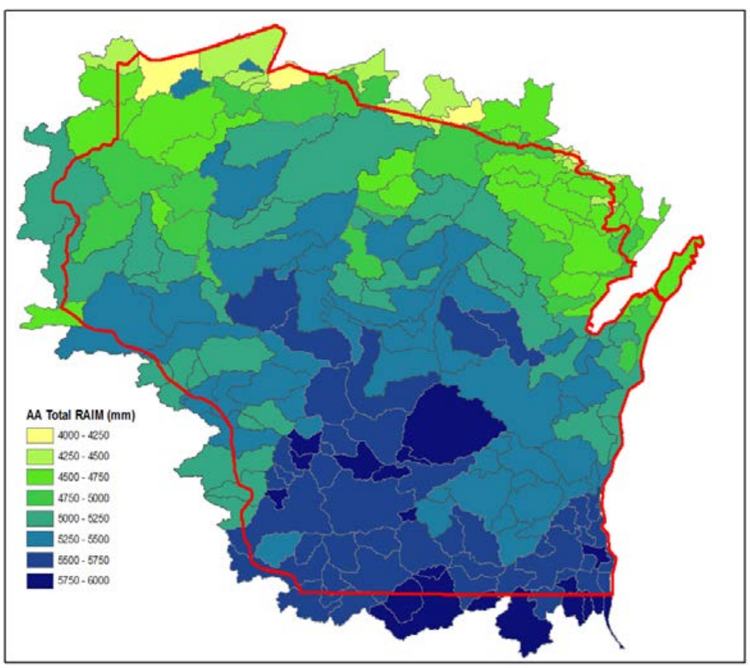


Figure 37. Analysis Accumulation of RAIM (mm) for 2011.

The highest values of AA RAIM were found over the southern half of Wisconsin aligning with the area of higher FMAP totals. Examining the presence of RAIM for all of the basins (Figure 40) indicated the median percent of time-steps where a value was greater than zero was 18%. The range was between 16 – 52%. Conversely, 82% of the time-steps for all basins generated no RAIM. Basins in north central Wisconsin had the highest incidence of RAIM presence which also coincides with the highest incidence of FMAP. It should be noted that one basin in particular generated RAIM much more often than the rest. The MRNM4 basin along the Michigan-Wisconsin border generated RAIM 52% of the time which was more than double the next highest RAIM incidence at 23.7%. Removing that basin produces a range for percent of time RAIM is produced from 16 – 24%.

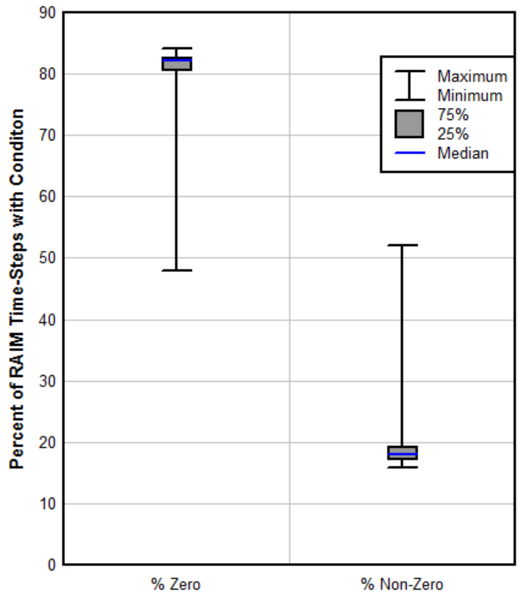


Figure 38. Distribution of percent of Analysis Accumulated time-steps where RAIM is equal to zero or greater than zero.

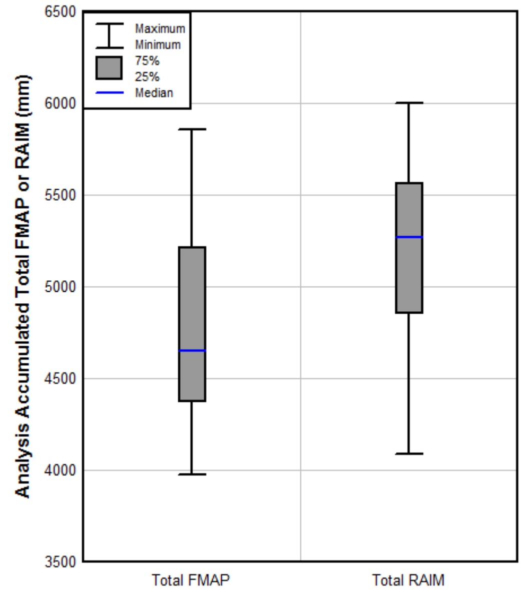


Figure 39. Distribution of Analysis Accumulated FMAP and RAIM for all 216 basins in Wisconsin.

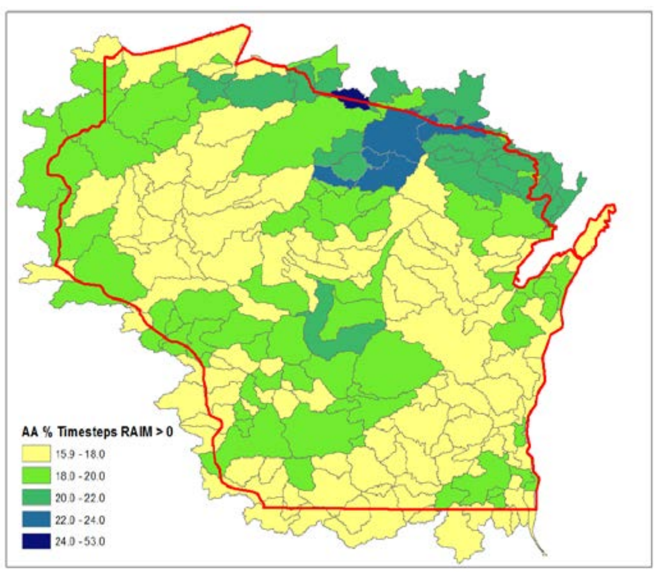


Figure 40. Percent of Analysis Accumulated time-steps with RAIM greater than zero for 2011.

4.2.3 SAC-SMA Interflow Runoff (INTRO)

The third model component and the most important is the interflow runoff generated from the SAC-SMA model. The importance is twofold in that interflow must be present for an event to occur and the accumulation of the simulated interflow for the event is used to stratify the risk for a particular basin. As computed for prior components, the AA median for all basins was found to be 484 mm with a range from 0.63 to 2,176 mm (Figure 41). The range, although rather large, is not entirely surprising as interflow is more dependent on basin model parameters and calibration than either FMAP or RAIM which can lead to differences in how runoff is generated.

The all basin median percent of time interflow was present was 20% and conversely 80% of the time no interflow was generated (Figure 42). The range for the percent of time interflow was present was between 0.1% and 50%. Spatially, the highest totals of AA interflow were located in central Wisconsin whereas the lowest totals were in basins in northern part of the state (Figure 43). Viewing Figure 43 also highlights some outlier basins that are not generating similar amounts of interflow when compared to their neighbors.

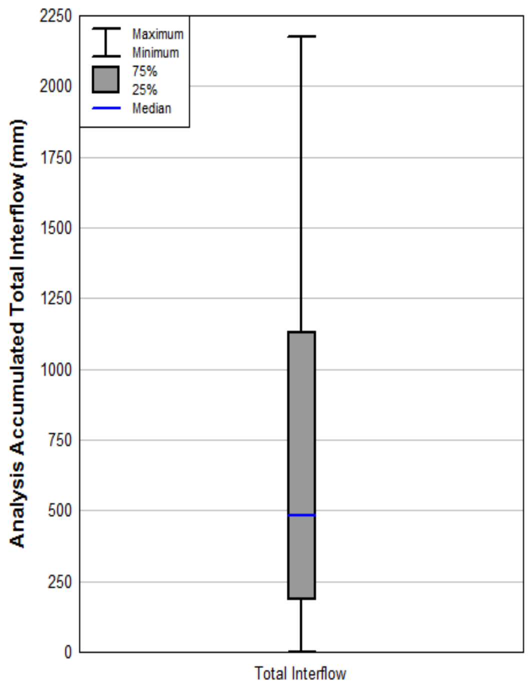


Figure 41. Distribution of Analysis Accumulated total interflow runoff (mm) for the 216 basins.

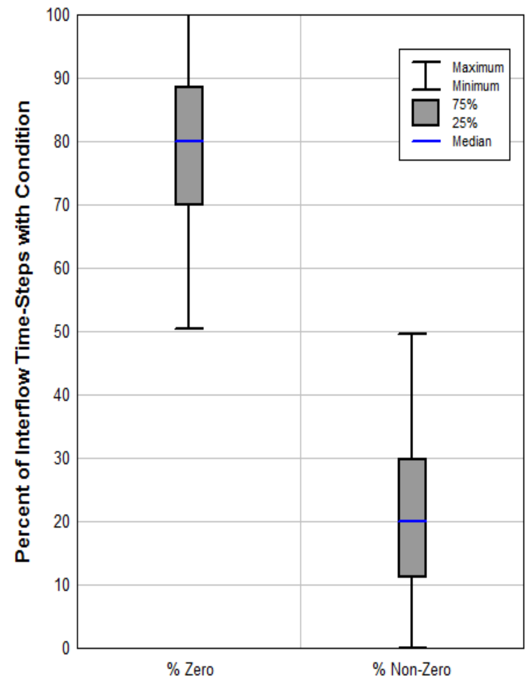


Figure 42. Distribution of percent of Analysis Accumulated time-steps where interflow runoff is equal to zero or greater than zero.

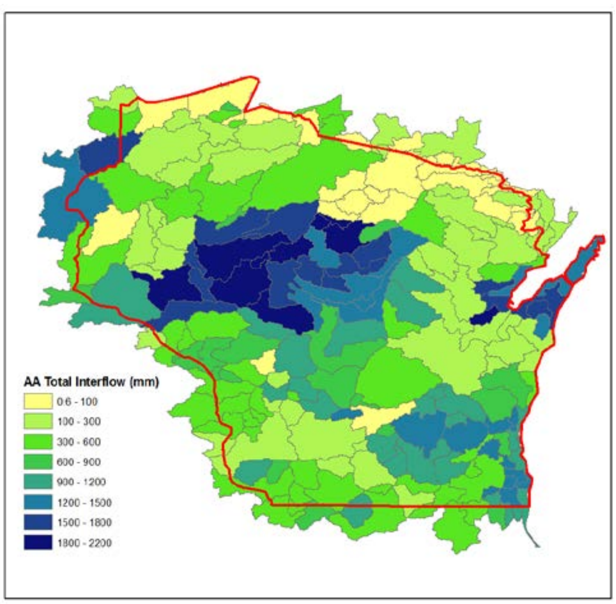


Figure 43. Spatial distribution of Analysis Accumulated interflow runoff for 2011.

The situation of relative high and low interflow producing basins next to each other does not necessarily insinuate a problem as long as the smaller producing basin produces interflow on a similar amount of time-steps. This is true for a couple reasons: (1) event runoff for a basin is specific to only that basin and not indicated to the users, and (2) basin specific thresholds should stratify the risk similarly for each of the basins. However, situations where a basin produces not only a small amount of interflow, but is also does not generate it very often can cause some consistency issues. In this situation, differences are likely due to calibrated parameters and no threshold manipulation can be accomplished to smooth out the RRAF map.

4.2.4 SAC-SMA Upper Zone Tension Water Deficit (UZTWD)

The final model component needed for simulated runoff events and evaluated in the 2011 forecast runs is the Upper Zone Tension Water Deficit (UZTWD) from the SAC-SMA model. Recall that this variable describes the fullness of the conceptual upper soil layer. This bucket has the ability to increase or decrease in value at every time-step. The summation of the time-step UZTWD values in this analysis has no physical meaning and instead is being used as an indicator of the ability for basins to build more or less deficits over time compared to neighboring basins. It is also worth mentioning that the SAC-SMA parameters are all connected, therefore the size of a basin's UZTW bucket can also impact how often it is saturated as well as how much interflow runoff can be generated.

Remembering that for a simulated runoff event to occur, the UZTWD must become equal to and remain at zero. The median percent of time-steps in the 2011 forecast analysis indicated that 17% of the time basins in Wisconsin the UZTWD was zero. The maximum amount of time for any basin being saturated was 43% while the minimum was 6% (Figure 44).

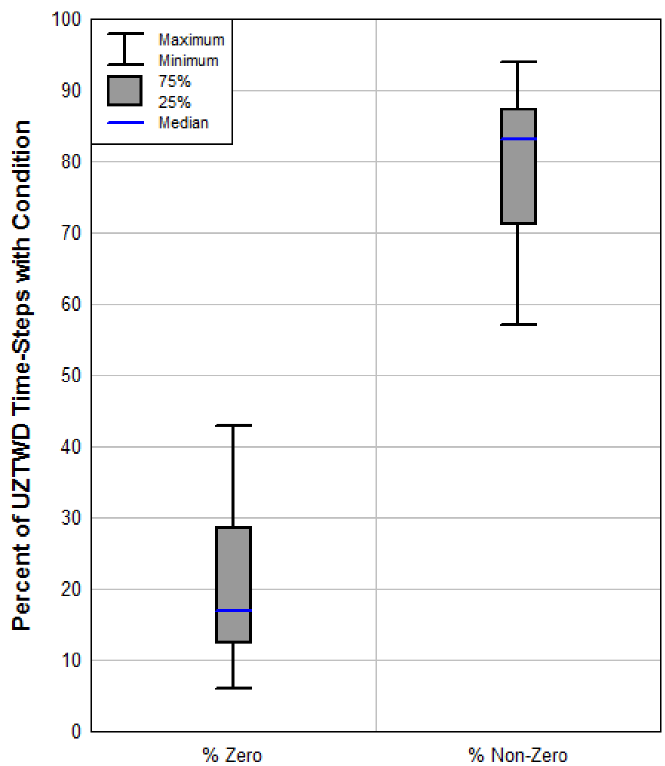


Figure 44. Distribution of percent of Analysis Accumulated time-steps where UZTWD equals zero or is greater than zero.

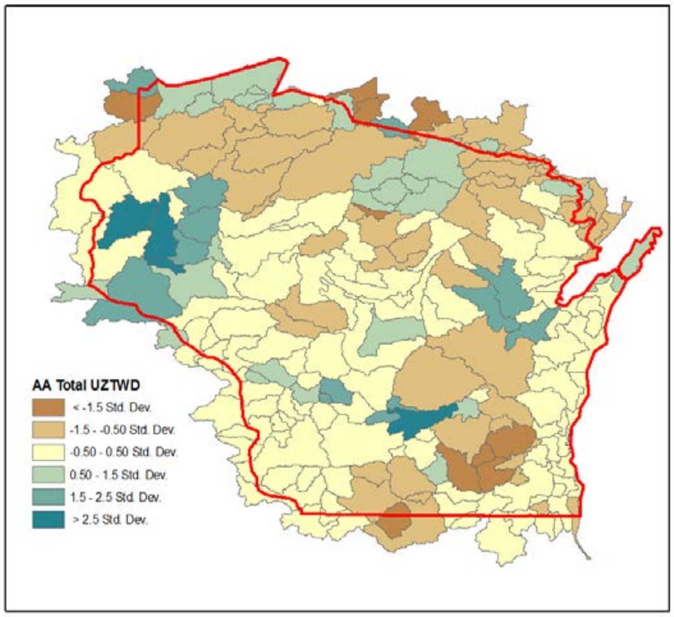


Figure 45. Spatial distribution of the Analysis Accumulation of UZTWD values. Areas of blue have built larger deficits over the year compared to lower values in the orange shades.

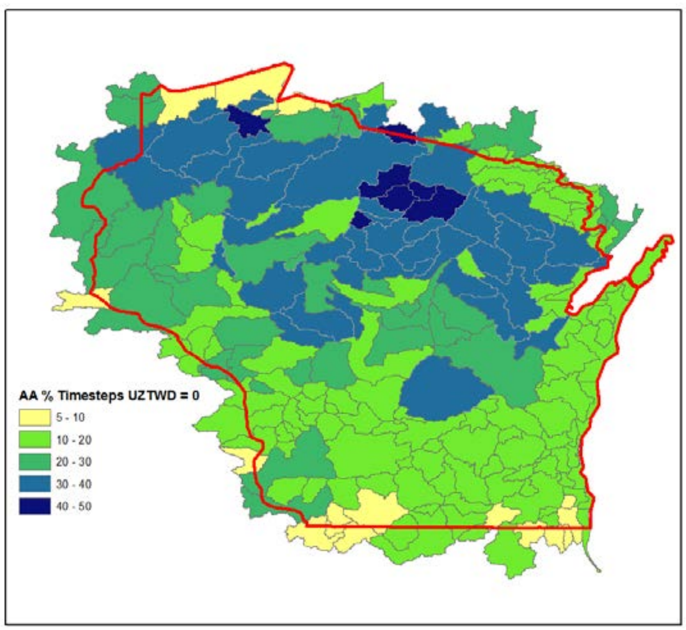


Figure 46. Spatial distribution of the percent of Analysis Accumulated time-steps where UZTWD was equal to zero.

The areas of highest accumulated UZTWD for 2011 were found to occur in western Wisconsin and a few other smaller pockets around the northern border of the state and an outlier basin (PDSW3) in south central Wisconsin (Figure 45). There were a few occurrences of much lower and much higher than normal accumulated UZTWD values found near each other. These could be red-flags for potential problems and were noted for further investigation. The percent of the time where the UZTWD was equal to zero is shown in Figure 46. Here the southern third of the state is generally in the 10-20% of the time whereas the northern half of the state was in the 30-40% category. Some basins with very low incidence of upper soil saturation are seen in this map which indicates those basins will be very unresponsive in the RRAF guidance. This finding indicated further model parameter evaluation of those basins was needed.

4.3 Analysis of Simulated Runoff Events

With the model component analysis complete, this section will summarize the AA simulated runoff events for 2011. The events will be broken down by quantity of occurrence and associated runoff, spatial distribution, and category definition. The amount of time in each category as well as the source of runoff will also be investigated.

4.3.1 Number of Simulated Events

In the 2011 Accumulated Analysis there were 40,389 simulated events from all 216 basins. The high risk category was simulated more often than moderate risk with CAT 3 scoring 22,820 events while CAT 2 only recorded 17,569 events. Essentially 57% of the events were CAT 3 and 43% were CAT 2 (Figure 47). However, breakdown by category did have quite a range with CAT 2 spanning from 19 – 89% of a particular basin's events while CAT 3 ranged from 11 – 81%. Evaluating the median behavior from the 216 basins provides a value of 198 events with 83 CAT 2 and 114 CAT3 (Figure 48). Total events for any given basin did range from 9 to 287 while CAT 2 ranged from 8 – 153 and CAT 3 ranged from only 1 to 176.

Examining the spatial distribution of total simulated events indicates that the basins with the highest occurrence of runoff events are in the southeastern part of the state as well as along the Door County Peninsula. The basins with the least amount of event activity appear on the northern shore of Wisconsin as well as in the marshy headwaters of the Wisconsin River in the north central part of the state (Figure 49). The distribution of basins on the high and low end of CAT 2 event generation was somewhat scattered across the state with several examples of basins in the top 10% and bottom 10% of occurrence next to one another (Figures 49 and 50). This spatial mixture can be explained generally due to basins in the far northern part of the state which generate very few events overall or generate very few CAT 3 events in comparison to the rest the basins.

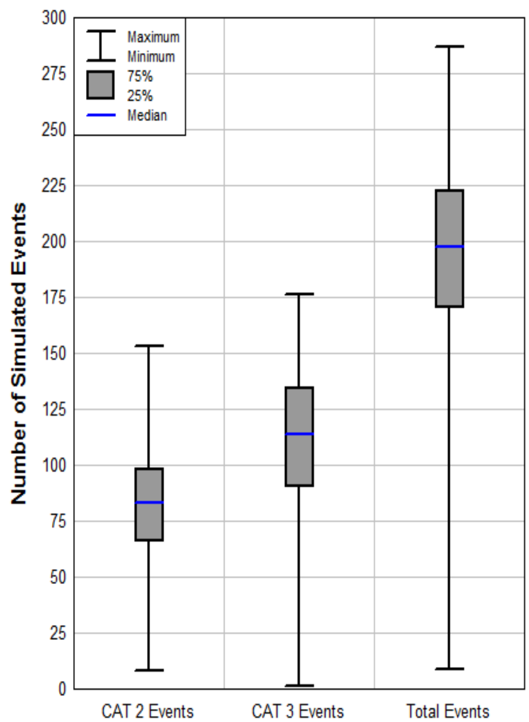


Figure 48. Distribution of Analysis Accumulated number of events in a basin for 2011.

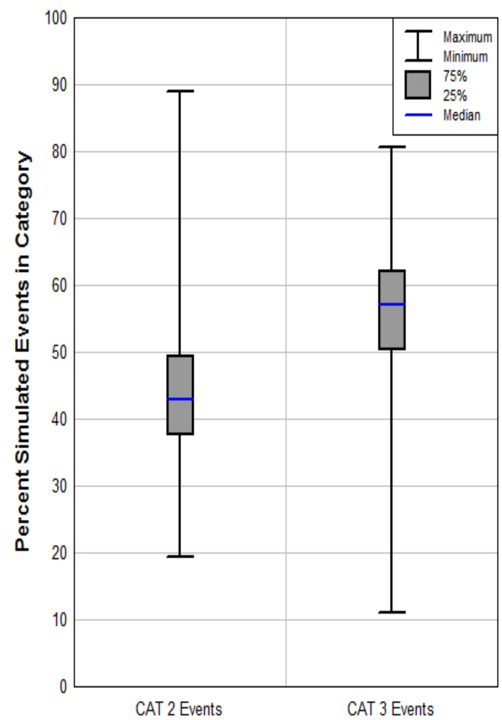
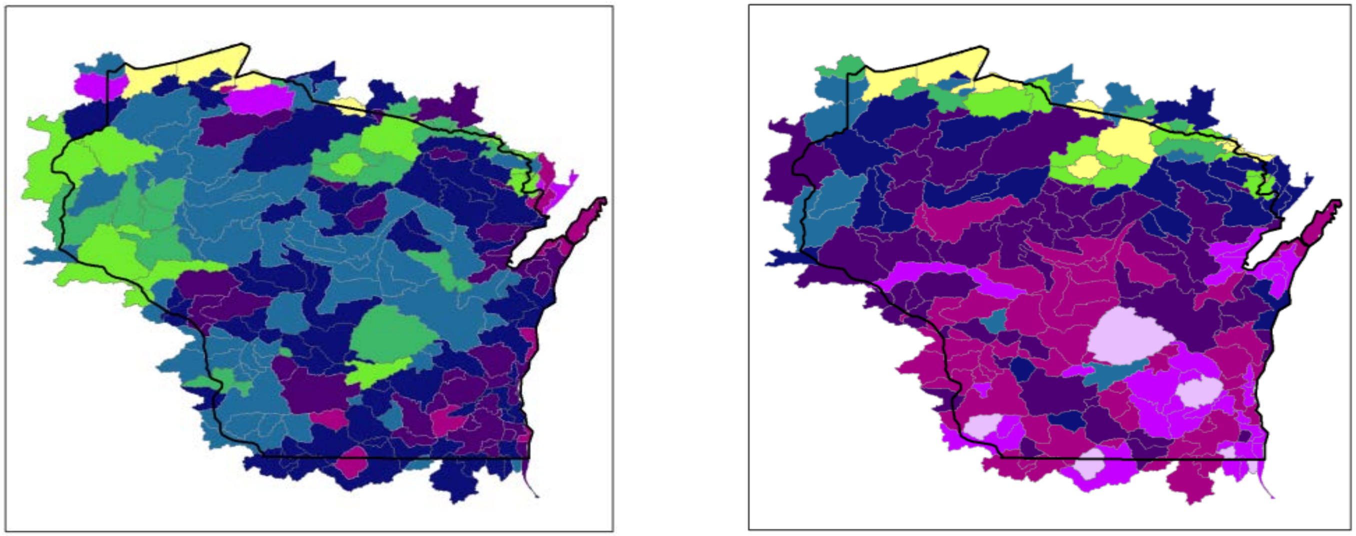
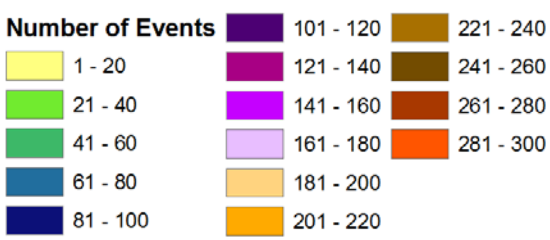


Figure 47. Distribution of the Analysis Accumulated percent of basin simulated events in each category for 2011.

The CAT 3 distribution was found to align very close to the total event pattern with lower incidence of the high risk events in the north and the highest generating basins for CAT 3 again in the southern third of the state. Figure 51 below presents the spatial distribution of the percent of basin events in each category that is summarized in Figure 48. The graphics shown support the findings described above with a northern bias for CAT 2 events and a southern bias for CAT 3 events. In addition, basins generally produce more CAT 3 events than the lower risk category.

As mentioned earlier when summarizing the model components, spatially evaluating the simulated runoff events, in total and by category, offers an excellent way to detect model parameter inconsistencies as well as highlighting areas that can be improved not only for the RRAF product but for overall streamflow forecasting as well.





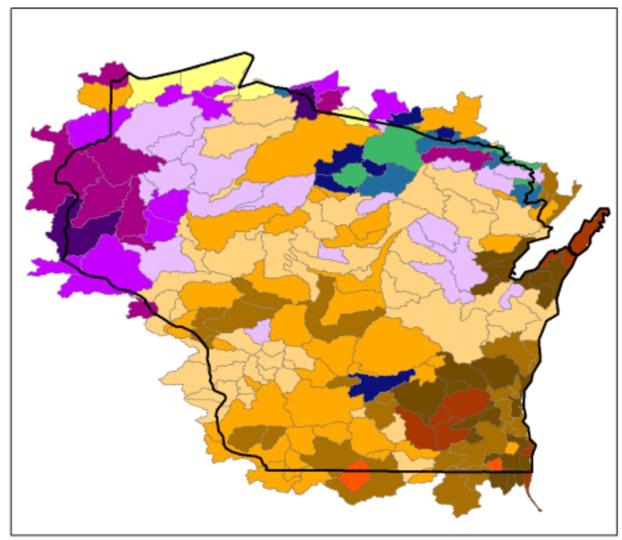


Figure 49. Spatial distribution of Analysis Accumulated CAT 2 events (upper left), CAT 3 events (upper right), and total events (lower right).

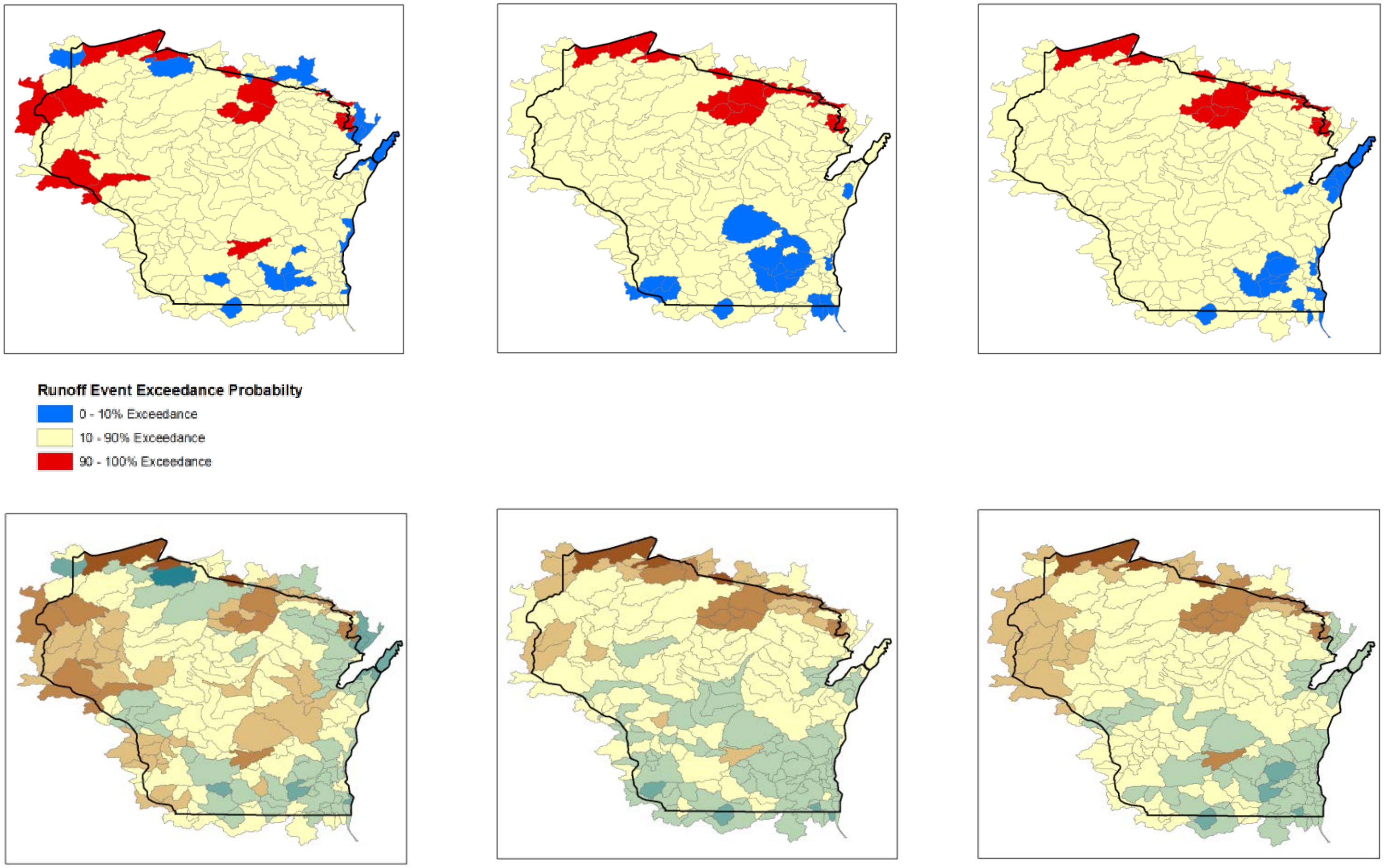


Figure 50. Spatial distribution of basins with the largest and smallest number of simulated runoff events in the 2011 Analysis Accumulation. CAT 2, CAT 3, and total number of runoff events are in the left, center, and right columns respectively. Top row is the top and bottom 10% of responding basins. Bottom row highlights the basins based on standard deviations.

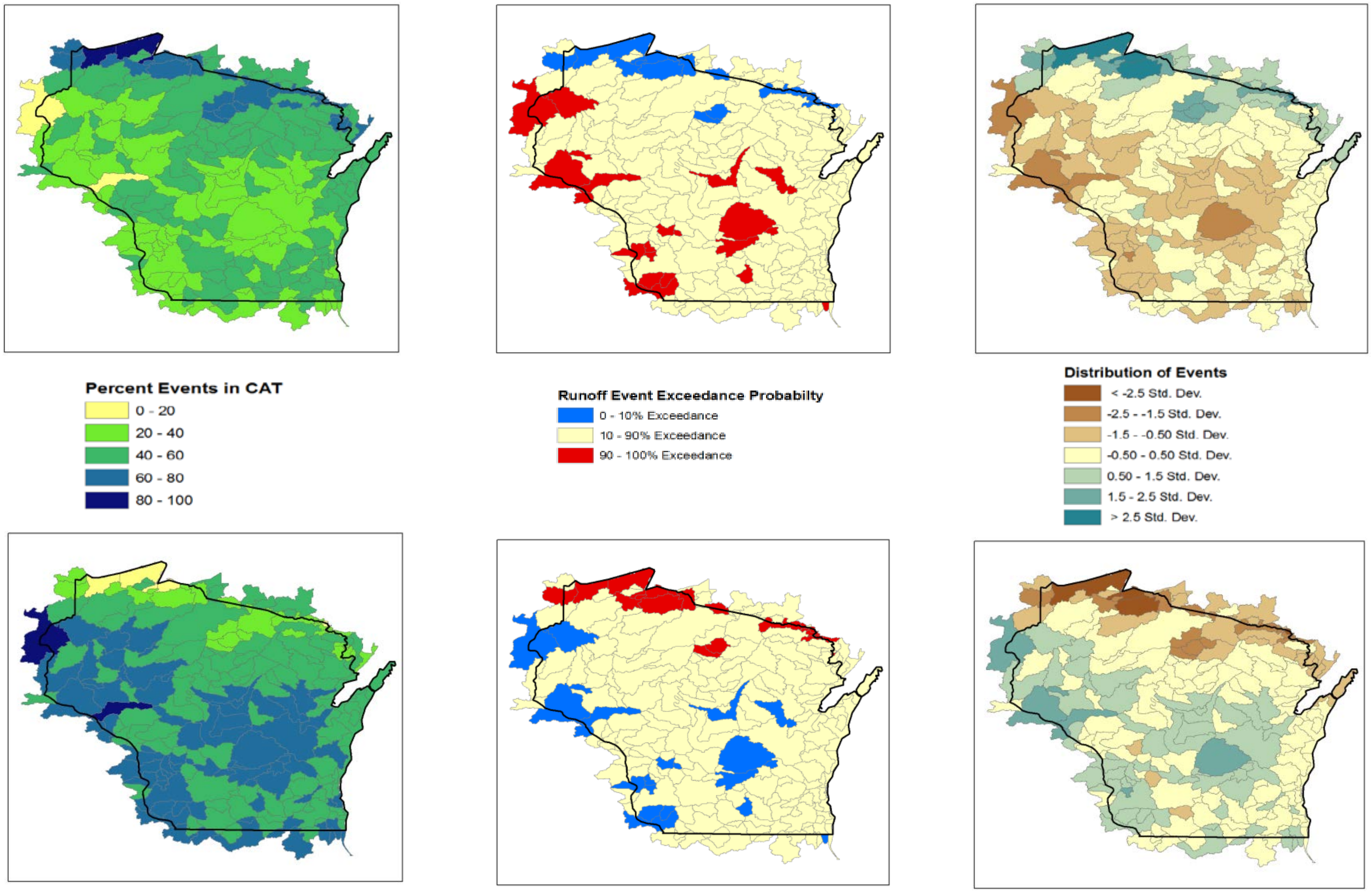


Figure 51. Spatial distribution of the percent of Analysis Accumulated simulated runoff events for a basin that were CAT 2 (top row) or CAT 3 (bottom row). Left column is the percent of all events for a given basin in each category. Center column is the top and bottom 10% basins based on event incidence in each category. Third column is the distribution of events in each category broken down by standard deviations.

4.3.2 Simulated Event Runoff

Now that the number of events has been summarized, the next step entails analyzing the runoff volumes generated by the events in the 2011 forecast analysis. The total runoff generated from all 216 basins in this analysis was 96,975 mm. Broken down by risk category, CAT 2 accounted for only 4,151 mm, or 4.3%, whereas the high risk events accounted for 92,824 mm, or 95.7% of the total volume generated.

On a basin basis, the Analysis Accumulated median CAT 2 runoff was only 15 mm with a range from 0.28 to 65 mm. On the other hand the CAT 3 median basin runoff was 358 mm with a range from 0.35 to 1,548 mm (Figure 52). The threshold process appears to be successfully parsing the larger events into the high risk category. Although CAT 3 events occur only 14% more often than CAT 2 events, 91% more runoff is associated with the CAT 3 events. Similar to the overall category runoff breakdown, the median basin values are 4.7% in CAT 2 runoff and 95.3% in CAT 3. The range for CAT 2 was 0.5 to 44.4% whereas the range for CAT 3 is 55.6 to 99.5% (Figure 53).

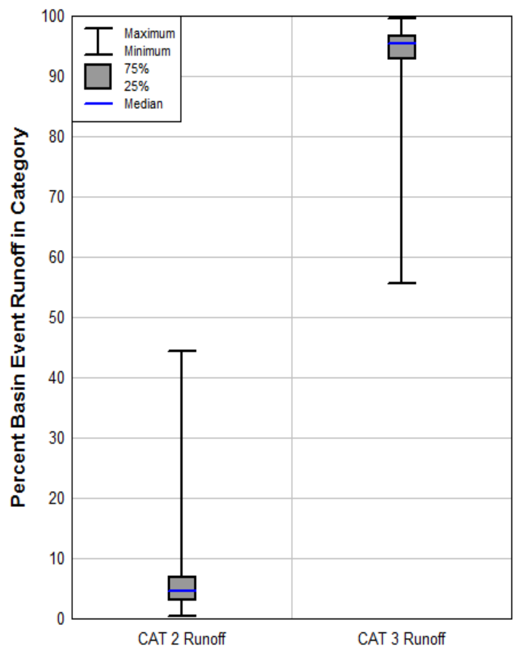


Figure 52. Distribution of the percent of event runoff in each category for a basin in the 2011 analysis.

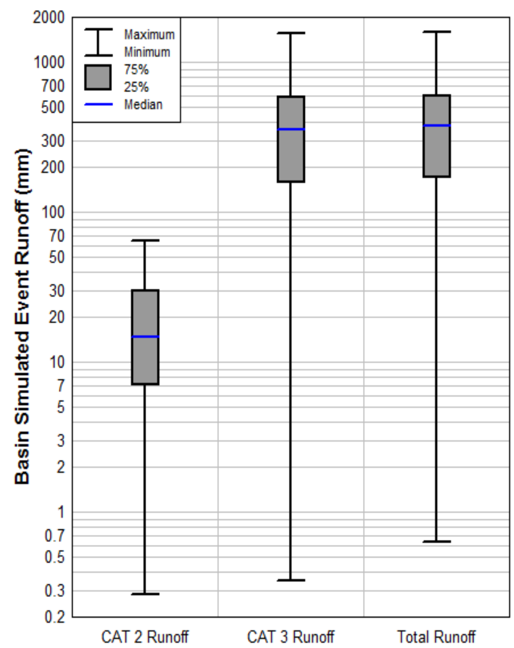


Figure 53. Distribution of basin event runoff for CAT 2, CAT 3, and all events in the 2011 analysis.

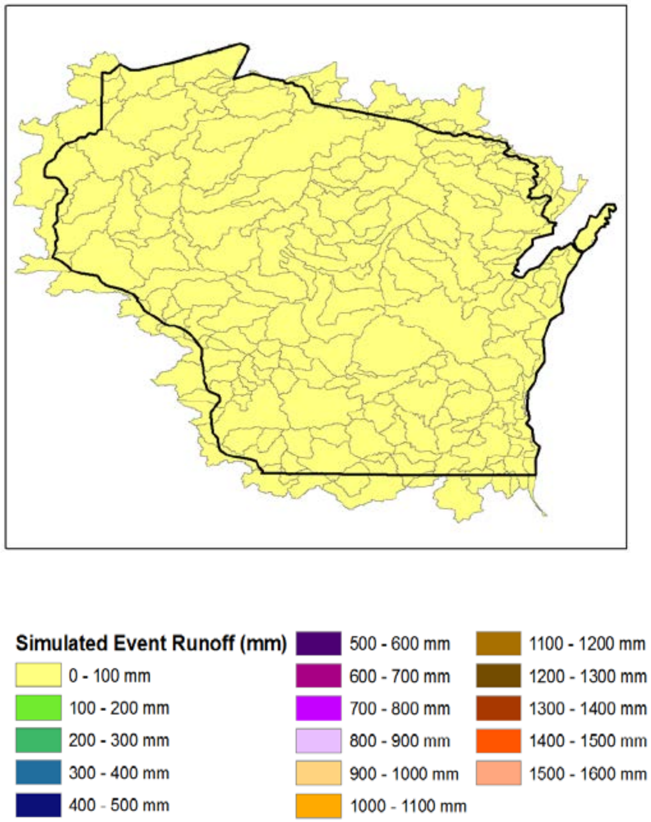
Spatially the pattern of AA event runoff (Figures 54 and 55) is very similar to interflow runoff (Figure 43) with the highest totals located in central Wisconsin. The occurrence of basins with dramatically different runoff totals near each other is not surprising just as it was not surprising in the interflow analysis. Variation in basin parameters across the state will induce some discrepancies in event runoff when compared to neighboring basins. Recall that event runoff totals are not a point of emphasis in the RRAF nor presented to the end users. Only the designation of an event and associated category are relayed to the user.

4.3.3 Percent of Time in each Category

Now that the number of events and quantity of runoff have been summarized for the 2011 analysis, the third aspect reviewed will be the amount of time the basins are slotted in each category. In addition to the CAT 2 and CAT 3 risk categories, a new category, CAT 1, will be used to designate the time when no runoff event has been simulated. Recall the 2011 analysis is derived from 365 forecast runs consisting of ten day windows. Therefore, for this analysis each basin has a maximum possible time allotment of 3,650 days or 87,600 hours.

The summation of every basin into the time categories indicated that 92% of the time no event was simulated (CAT 1) (Figure 56). Moderate risk conditions were simulated 2% of the time and CAT 3 was simulated 6% of the time. The presence of any type of runoff event is therefore 8% of the analysis period (Figure 57), not to be confused with 8% of the calendar year. On a basin median value approach, the breakdown in categories was essentially the same. The range for CAT 1 was between 88.5% and 99.9%. The minimum for CAT 2 was 0.1% whereas the maximum percentage of time was only 7.1%. For CAT 3 the spread was between 0.01% and 9.6%. And finally, for any type of runoff event the percent of time covered was between 0 .1% and 11.5%.

Spatially, the basins in Wisconsin with the highest percentage of time with no runoff events simulated were along the northern edge of the state and in the headwaters of the Wisconsin River (Figure 58). In addition, an outlier basin near Prairie Du Sac (PDSW3) in south central Wisconsin is easily seen. As indicated by the statistics provided above, CAT 3 has higher percentages of occurrence compared to CAT 2 in most basins except for a few which are again located in the far northern portions of the state.



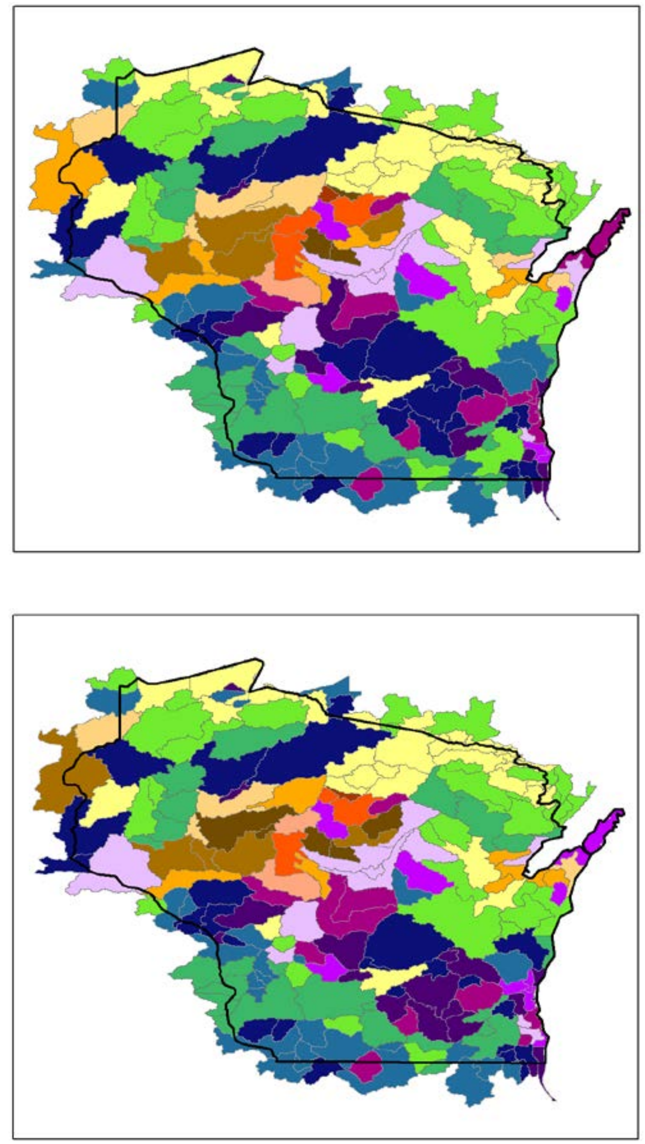


Figure 54. Spatial distribution of CAT 2 (upper left), CAT 3 (upper right), and total event runoff (lower right) in mm for the 2011 forecast analysis.

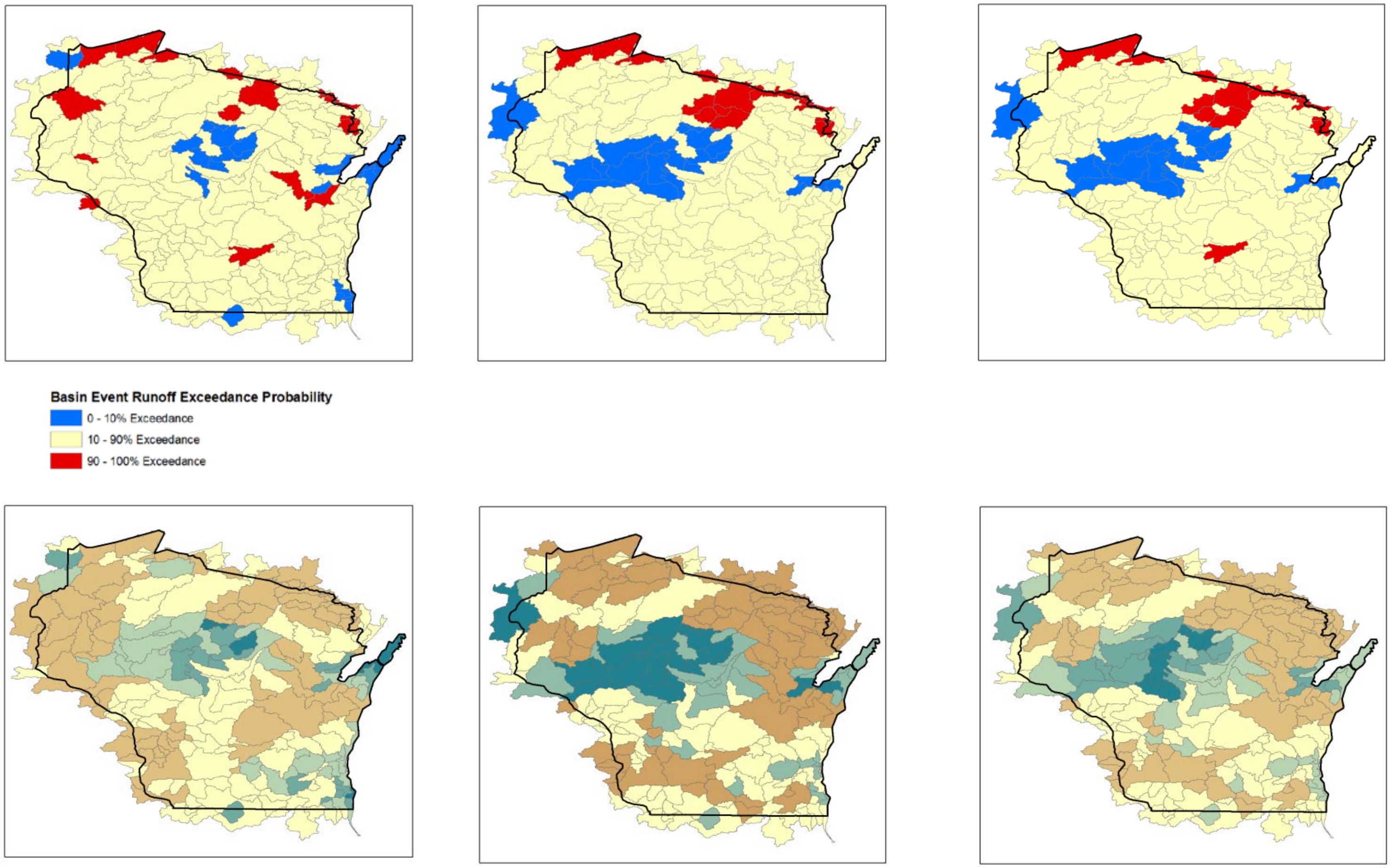


Figure 55. Spatial distribution of basins with the largest and smallest event runoff in the 2011 Analysis Accumulation. CAT 2, CAT 3, and total event runoff (mm) are in the left, center, and right columns respectively. Top row is the top and bottom 10% of responding basins. Bottom row highlights the basins based on standard deviations.

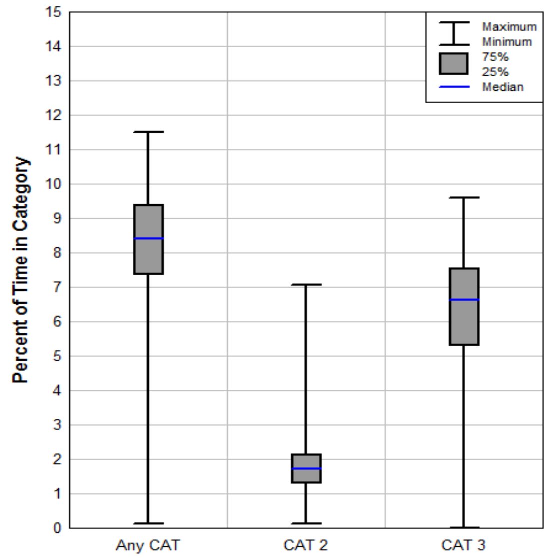


Figure 56. Distribution of the percent of total 2011 analysis period where any events, CAT 2, or CAT3 events are simulated for a basin.

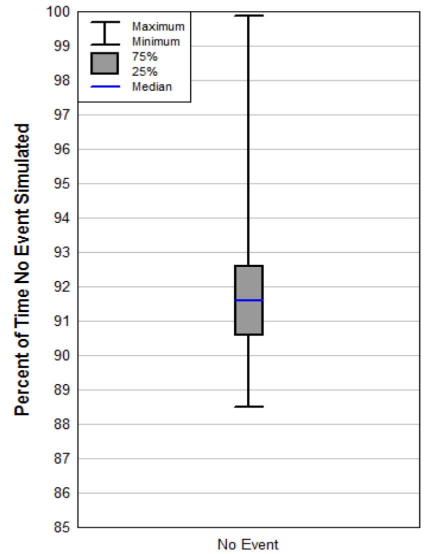


Figure 57. Distribution of percent of the total 2011 analysis period where no runoff event is simulated (CAT 1).

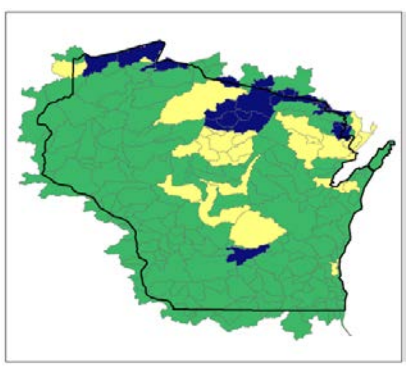
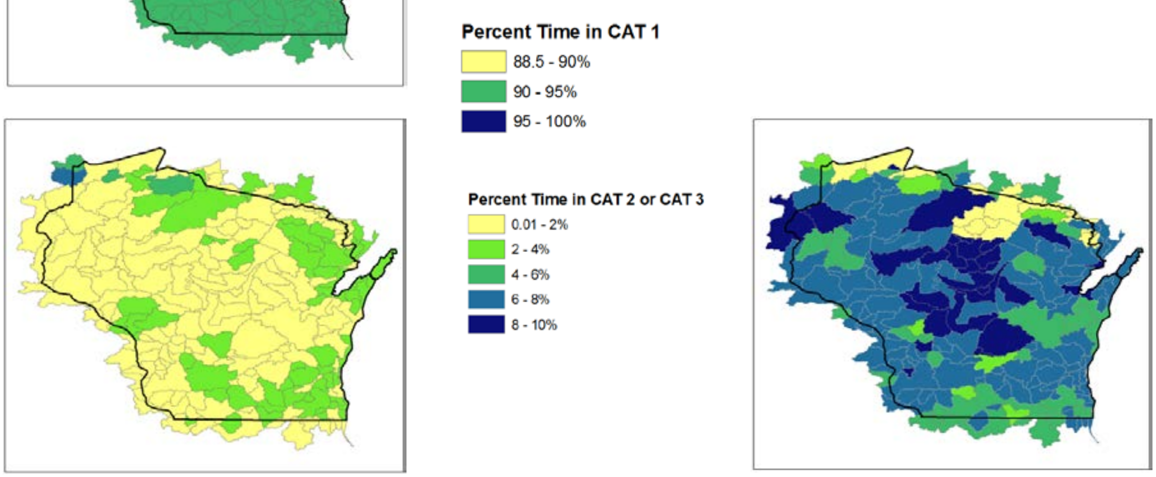


Figure 58. Percent of Analysis Accumulated time each basin was in the three categories. CAT 1 (no event) is left, CAT 2 is bottom left corner, and CAT 3 is bottom right corner.



So far it has been found that there are more CAT 3 events for a basin and they are associated with much more runoff when compared to CAT 2 events. This section compounded that trend by indicating the amount of time CAT 3 events are simulated is also more than that for CAT 2 events. Further refining the time breakdown can be accomplished by only considering time when any event is simulated for a basin. In a median basin value perspective, 77% of the time when a basin simulates an event during the analysis period it was defined as CAT 3. Therefore, the remaining 23% of the time those events were CAT 2.

4.3.4 Analysis Accumulated Values Compared to Historical Simulation

At this point an interesting evaluation can be made between the 2011 Analysis Accumulated values described above versus the historical summary statistics introduced earlier in section 3.5. Unfortunately it should be stressed that this is not a direct comparison due to the differences in the two data sets. The historical run is a continuous simulation over the specified time span fed with quality controlled observed mean forcing data only. On the other hand, recall that the 2011 analysis is strictly forecast data for a 10-day window and by default is injecting varying forecast precipitation and temperature estimates into the dataset. Therefore, any given day will be simulated up to 10 times with varying forecast forcing data and included in the final basin total.

Although not available at this time, in the future a time-series from the operational model runs will be created that will be analogous to the historical record. The plan is to save a runoff event record for each day for each basin based on observed forcing data in the operational model. That time-series summed over the year will provide a closer representation of what the model perceives as reality where the forecast component is not included. None the less, comparing the available historical simulation versus the 2011 analysis can still identify some interesting trends on the model behavior.

Interestingly, the comparison shows that the two datasets produced very similar results (Table 20). The 2011 forecast analysis hedged towards the CAT 3 category for number of events, quantity of runoff, and time in category. This is not that surprising given that the 2011 analysis includes forecast precipitation which often times is higher in magnitude compared to the actual observed value. In addition, the increased opportunity for events during future runoff conditions could also skew the 2011 analysis values to a higher overall percentage.

Table 20. Comparing event runoff characteristics between the historical simulations used during validation and 2011 forecast analysis.

Comparison of Median Basin Runoff Event Summary Statistics for Historical Simulation and 2011 Forecast Analysis.

	Historical Simulation	2011 Analysis
Percent Events CAT 2	60%	57%
Percent Events CAT 3	40%	43%
Percent Event Runoff CAT 2	8%	5%
Percent Event Runoff CAT 3	92%	95%
Percent Time No Event	88%	92%
Percent Time CAT 2	5%	2%
Percent Time CAT 3	7%	6%
Percent Time Any Event	12%	8%

4.3.5 Simulated Events Boolean Perspective

For a more general review on basin simulated event activity a Boolean analysis was completed on the 2011 forecast runs. In this scenario each basin for every daily run was tested to determine if a simulated event was generated at any time in the 10 day window. If there was at least one event simulated the basin was assigned a “1”, otherwise it was given a “0”, resulting in a maximum possible score of 365. Specifically, the number of daily runs with CAT 2, CAT 3, any events, and no events were tallied for the basins. Further, daily runs were exclusively typed with only one category where higher risk trumped lower risk. For example, if a basin had at least one CAT 2 event and at least one CAT 3 event, it was only categorized as a CAT 3 for that day.

The sum of all the daily runs for one year for 216 basins is 78,840. The analysis showed 30,613 basin-daily runs (39%) with any event simulated and 48,227 basin-daily runs (61%) with no event simulated. The median basin tallies indicated 151 daily runs with any event (41%) and 214 daily runs with no event (59%). The range for runs with any event spanned from a max of 218 (60%) to a minimum of only 9 (2%). Conversely, the number of daily runs for a basin with no event ranged from 147 to 356 (40 – 98%). Out of the median value of 151 daily runs with an event, 63 (42%) were CAT 2 and 88 (58%) were CAT 3 (Figures 59 and 60). Compared to the total possible number of runs, CAT 2 was present in 17% whereas CAT 3 was detected in 24% of the runs. The number of daily runs with an event ranged between 8 – 109 (2 – 30%) for CAT 2 and 1 – 128 (0.3 – 35%) for CAT 3.

Spatially the distribution of the Boolean test on the daily runs with at least one runoff event is very similar to the total number of events by category evaluated earlier in section 4.3.1. The more active basins were found in the southern part of Wisconsin while the northern areas of the state, and to some extent western Wisconsin, recorded fewer days with any type of runoff event (Figures 61 – 63).

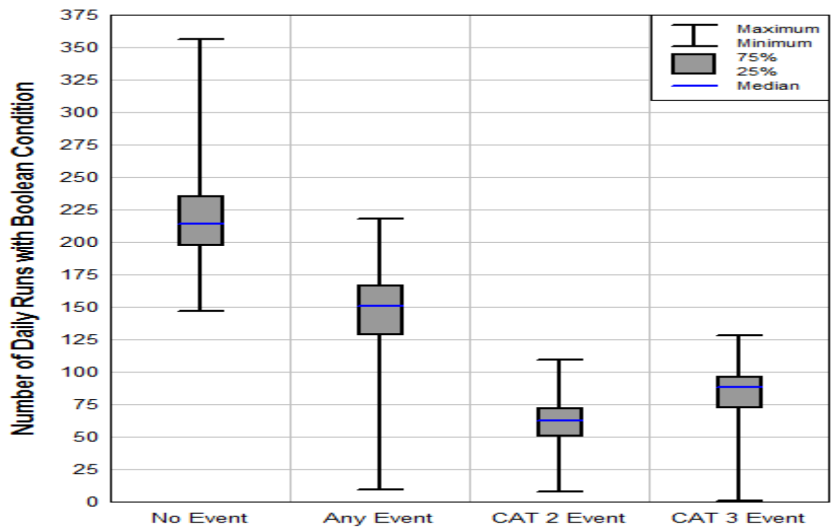


Figure 59. Analysis Accumulated distribution of number of daily forecast runs tested with Boolean condition for no event, any events, or either risk category.

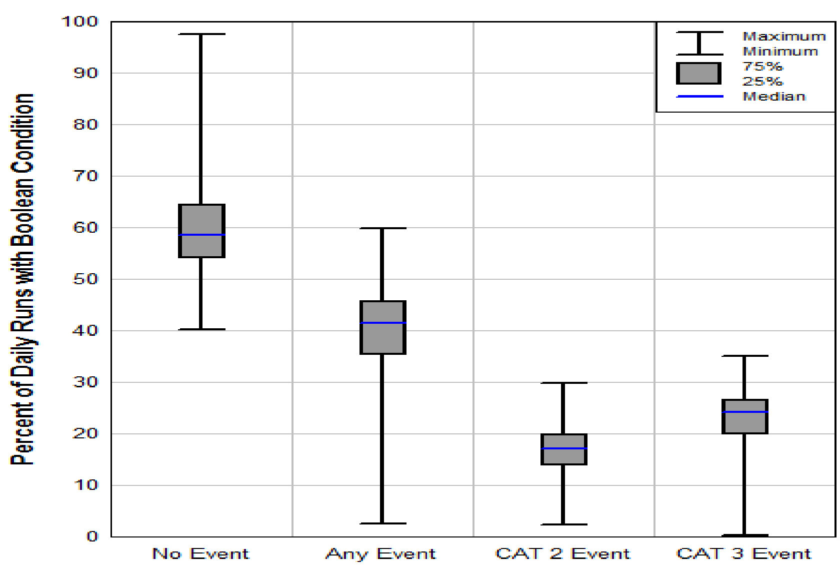
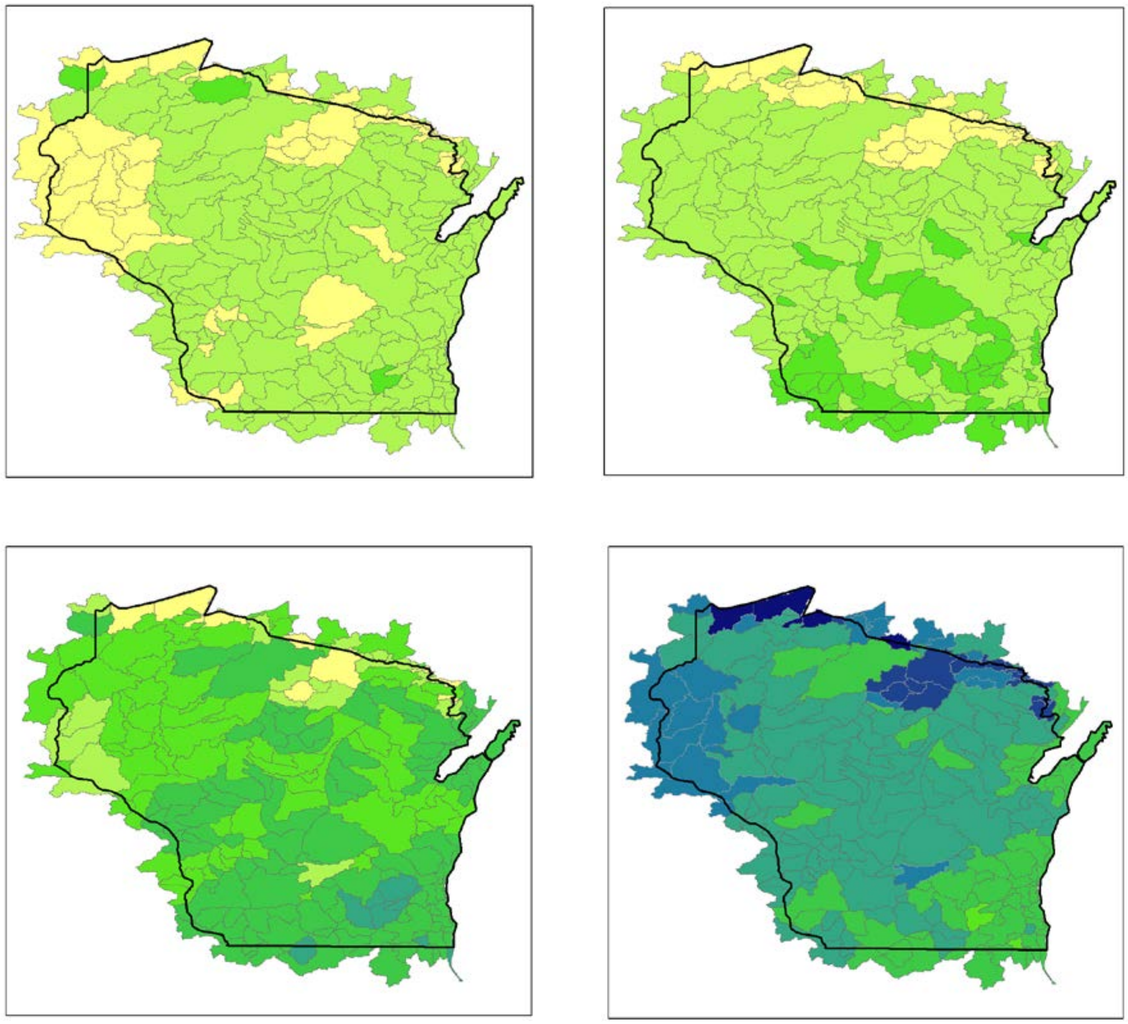


Figure 60. Analysis Accumulated distribution of percent of daily forecast runs tested with Boolean condition for no event, any events, or either risk category.



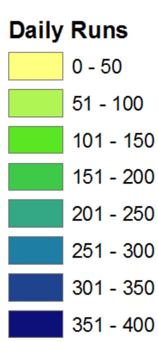


Figure 61. Spatial distribution of Analysis Accumulated Boolean daily run totals for the basins in Wisconsin. Daily runs were tallied for each basin if at least one event was present (or not) in the 10 day forecast window. Maximum score would be 365. Upper left is CAT 2, upper right is CAT 3, lower left is any event, lower right is no event.

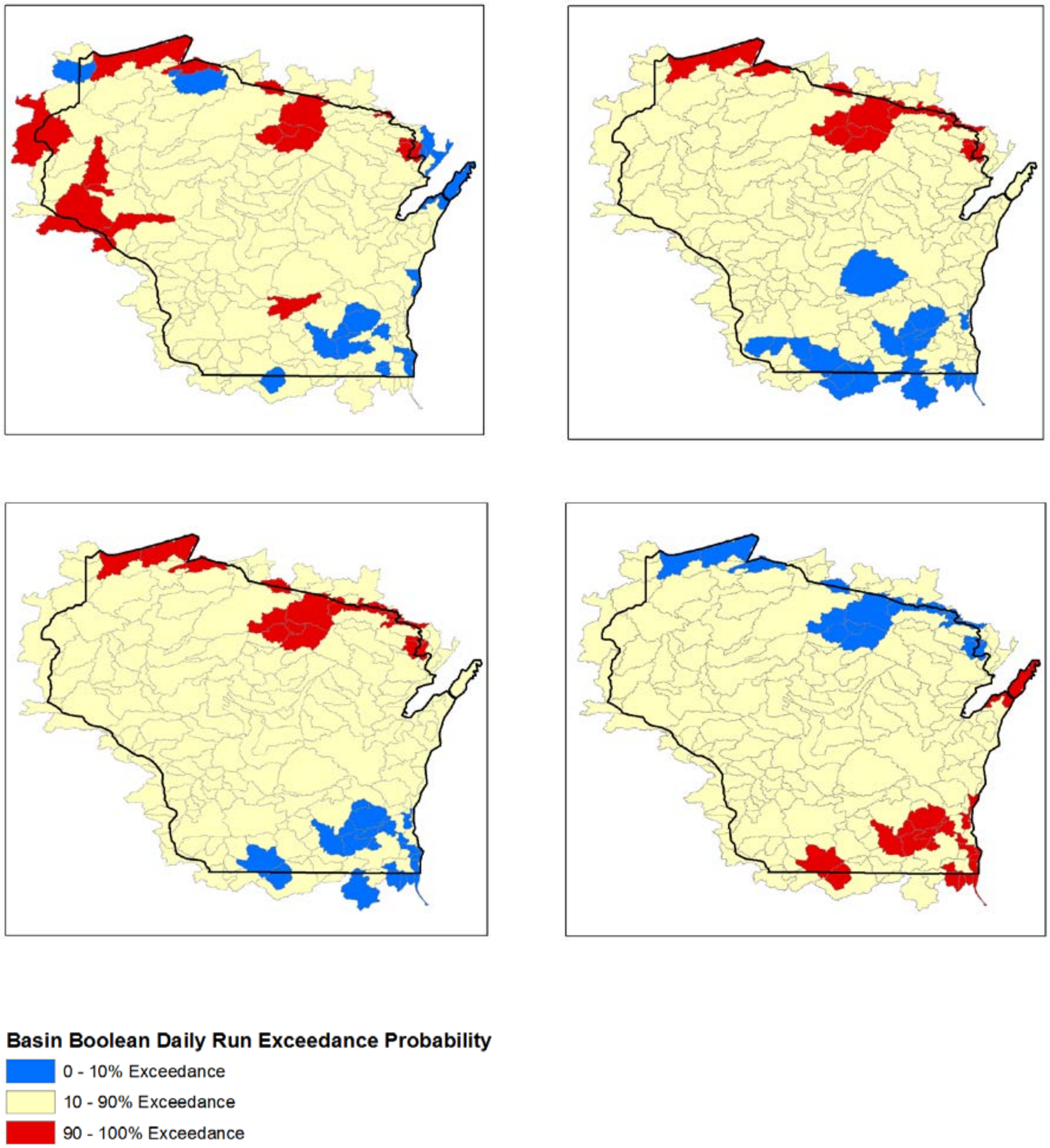


Figure 62. Spatial distribution of Analysis Accumulated Boolean daily run totals stratified by the most and least active basins in Wisconsin. Daily runs were tallied for each basin if at least one event was present (or not) in the 10 day forecast window. Upper left is CAT 2, upper right is CAT 3, lower left is any event, lower right is no event.

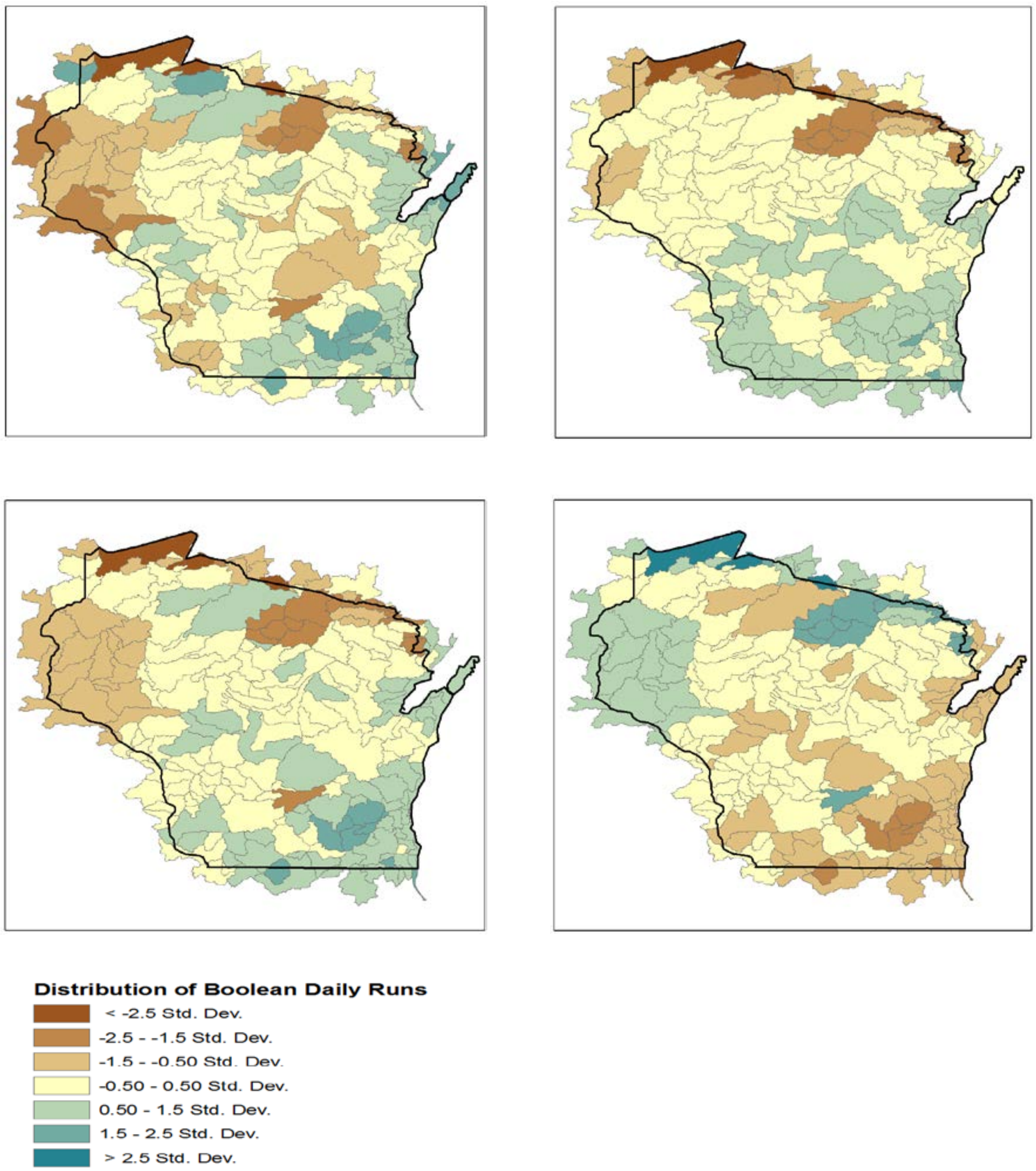


Figure 63. Spatial distribution of the basins with more or less daily runs containing runoff events in Boolean perspective compared to overall average. Top left is CAT 2, top right is CAT 3, bottom left is any event, and bottom right is no event.

4.3.6 Introduction of Warning Day Perspective

This section will review perhaps the most meaningful summary of the 2011 RRAF forecast analysis as it closely replicates how the RRAF runoff events are processed and portrayed to the users on the DATCP webpage. Incorporating a Boolean approach and forecast windows required the definition of "Warning Days" (WD). The forecast windows consist of three 72-hour time frames over the first five days and one 120-hour time frame for the second five days of the forecast run. The Boolean test is applied to each of the four Warning Days independently. A Warning Day is tallied in a risk category if at least one runoff event is simulated to occur during its time span. Below are the time spans for each Warning Day with the hours from start of the daily model run indicated by the subscript (Figure 64). Figure 65 reiterates why the Analysis Accumulations referred to in Chapter 4 are not equivalent to annual totals given that any particular calendar day can be simulated and tallied up to ten times.

- WD 1: $T_{00} - T_{72}$ (Forecast Days 1 – 3)
- WD 2: $T_{24} - T_{96}$ (Forecast Days 2 – 4)
- WD 3: $T_{48} - T_{120}$ (Forecast Days 3 – 5)
- WDX: $T_{120} - T_{240}$ (Forecast Days 6 – 10)

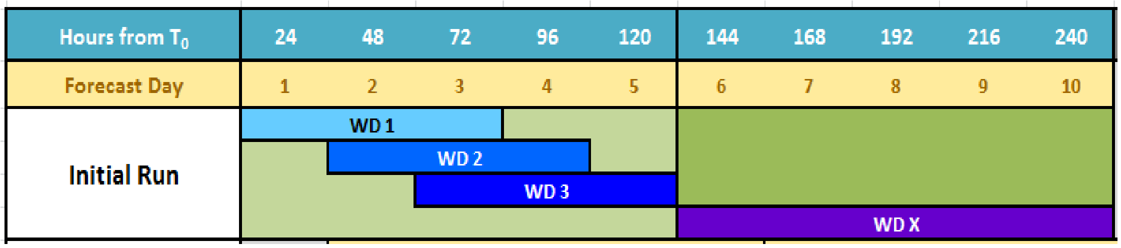


Figure 64. Schematic of the Warning Day concept which breaks the 10-day forecast run into three 72-hour windows and one five day window.

This analysis evaluates CAT 2 and 3 events for each Warning Day separately so that a given Warning Day can be counted as a hit for both categories if both type of runoff events occurred during the time frame. In addition a tally for Warning Days where any event occurs is recorded for the basins. On the RRAF webpage, basins are coded based on the highest risk simulated. Therefore if a Warning Day has both CAT 2 and CAT 3 events simulated, the basin will be coded red for the more threatening risk.

The Warning Day approach however is not exactly similar to the webpage behavior over the year. During late spring through late fall the webpage only uses Warning Days 1 – 3 where three risk categories are used. Note that during this time WDX is not used. During late fall and early winter the RRAF Working Group makes a decision to activate "Winter Mode" when frozen soils and/or snow cover becomes the majority condition in the state. In Winter Mode, only two categories are used (no event = winter risk and any event = high risk). In addition the forecast window expands to the end of the 10-day run. For example, in Winter Mode, Day 1 will be a Boolean test for any event in the next 10 days. Day 2 will be a Boolean test for any event in the remaining 9 days of the forecast run.

As the Winter Mode activation and deactivation can change with the year, it is not included in this Warning Day analysis. Instead, the entire year is summarized as if it were counted in the normal mode with low, moderate, high risk categories for the four Warning Day windows.

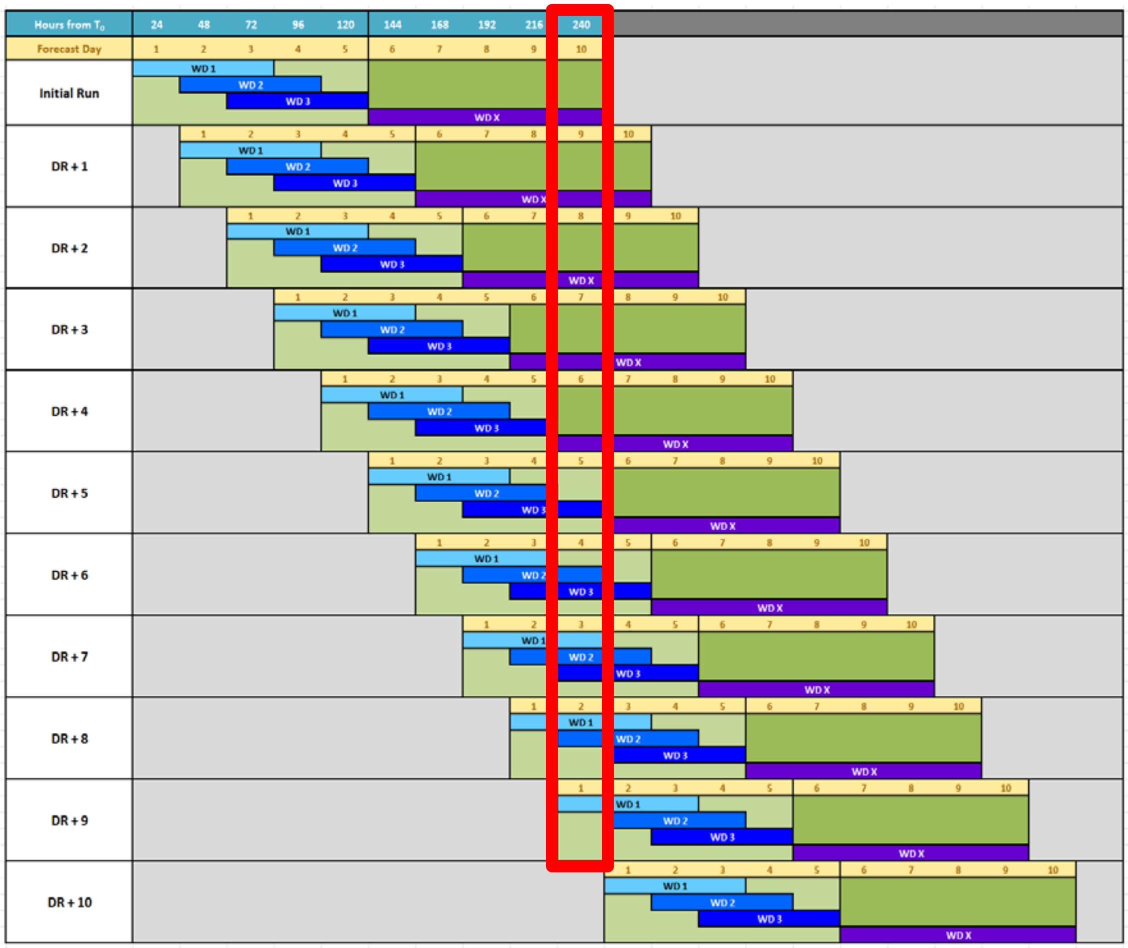


Figure 65. Example of cascading daily forecast runs which highlights how a particular day in 2011 could be simulated and tallied up to 10 times in the analysis.

Why 72-hour windows for the Warning Days? Recall that Chapter 2 referenced regulations requiring incorporation of manure into the soil within 3 days of application. Desiring to combine the forecast capabilities of the NCRFC models with these regulations led DATCP and the RRAF Working Group to proceed with that time window. A scenario illustrating the reasoning for this decision could include a producer planning to apply manure on a pleasant Monday morning and incorporating later on Wednesday. However, perhaps late Tuesday night a frontal system is expected to bring rain to the area or, less obviously, perhaps strong, warm, southerly winds will be developing which will help melt the snowpack before the producer can return to the fields. In these scenarios the manure is exposed to potential runoff conditions before the farmer could incorporate. Further, even if incorporation was accomplished, the possibility that manure and nutrients could still be leached from the fields into nearby water bodies due to intense future runoff conditions suggests taking into account the next couple days and waiting to apply was the better choice.

In this scenario, the producer would have checked the RRAF webpage on the nice and sunny Monday and saw moderate or high risk for that day because within 72 hours a runoff event was expected. Clicking on the forecast for Wednesday (Day 3) the producer could have saw low risk expected suggesting an alteration in his schedule would be the best idea to limit the potential for causing contaminated runoff from his fields while maximizing the use of his manure resources.

4.3.7 Warning Day Analysis of Simulated Events

The frequency of forecast runs with a simulated runoff event present across the various Warning Day periods is similar to the trend seen earlier in section 4.2.1 for daily forecast precipitation. The analysis accumulated FMAP summary indicated forecast days 2 and 3 had the highest totals. Therefore it is hardly a surprise to learn that Warning Days 2 and 3 are also the leaders in occurrence of runoff events on a basin median basis. This trend was seen when evaluating CAT 2 and CAT 3 separately as well as when any event was investigated (Table 21).

The basin median percent of daily runs where WD 2 and 3 had CAT 2 occurrence was 15% and 12% respectively. For CAT 3 the values were a bit higher at 23% and 20% for moderate and high risk events respectively (Table 22). Further, when evaluating whether any type of event occurred percentages of 32% and 29% were scored for WD 2 and WD 3.

Table 21. Daily forecast runs where a basin recorded at least one runoff event in a given Warning Day for CAT 2, CAT 3, and any event in 2011.

	WD 1	WD 2	WD 3	WD X
CAT 2 Median	**28**	**55**	**45**	**24**
CAT 2 Max	54	100	83	51
CAT 2 Min	5	6	4	0
CAT 3 Median	**36**	**82**	**74**	**41**
CAT 3 Max	64	118	104	58
CAT 3 Min	1	1	1	0
Any Median	**64**	**118**	**104**	**59**
Any Max	104	171	145	77
Any Min	7	7	5	0

Table 22. Percent of total daily forecast runs where a basin recorded at least one event in a given Warning Day for CAT 2, CAT 3, and any event in 2011.

	WD 1	WD 2	WD 3	WD X
CAT 2 Median	**8%**	**15%**	**12%**	**7%**
CAT 2 Max	15%	27%	23%	14%
CAT 2 Min	1%	2%	1%	0%
CAT 3 Median	**10%**	**23%**	**20%**	**11%**
CAT 3 Max	18%	32%	29%	16%
CAT 3 Min	0.3%	0.3%	0.3%	0%
Any Median	**17%**	**32%**	**29%**	**16%**
Any Max	29%	47%	40%	21%
Any Min	2%	2%	1%	0%

It is important to remember the story these Warning Day occurrence percentages are telling. These values are not equivalent to the base metrics discussed earlier in sections 4.3.1 - 4.3.4 which described the quantity of and characteristics of runoff events in 2011 forecast runs. Instead, and perhaps more importantly, these values are similar to the behavior the RRAF webpage over the year portrays to the manure producers in Wisconsin. Being Boolean based, Warning Days do not correspond to quantity or duration of runoff events, but instead relate to how many mornings a producer will notice his or her basin is not suitable for manure application based on the model forecast.

For almost a third of the daily runs in 2011, the RRAF indicated any type of runoff event was expected for the typical basin on Warning Days 2 and 3 (Figures 66 and 67). Further, around 20% of the daily runs included forecast conditions for high risk on those same two Warning Days. At first glance these values seem alarming, however, it is important to focus on what the Warning Days entail. They are not saying 20 – 30% of the total simulated time runoff events are being simulated. Those values were much lower and discussed in section 4.3.3. Requiring just the presence of an event overlapping a 72-hour window suggests that one forecast runoff event could activate more than one Warning Day period for several consecutive daily runs.

The results of the Warning Day analysis, combined with earlier knowledge about false alarm rates, do suggest some concern that users are being warned by the RRAF too often. Further evidence for this concern is obtained by reviewing the maximum basin Warning Day occurrences found in Table 22. For WD 2 and 3, these indicate there is a basin that is forecast to have a runoff event occurring on 40 – 47% of the daily runs while almost a third of a time a basin is indicating high risk. On the other extreme, basins exist which hardly ever have runoff risk forecast which is also a concern.

The data presented in Figure 67 is shown spatially in Figure 68 below. As with many images shown earlier, the southern portion of the state indicates the highest percentages of daily Warning Days with events. It is also clear that WD 2 and WD 3 are where most of the runoff event activity is shown.

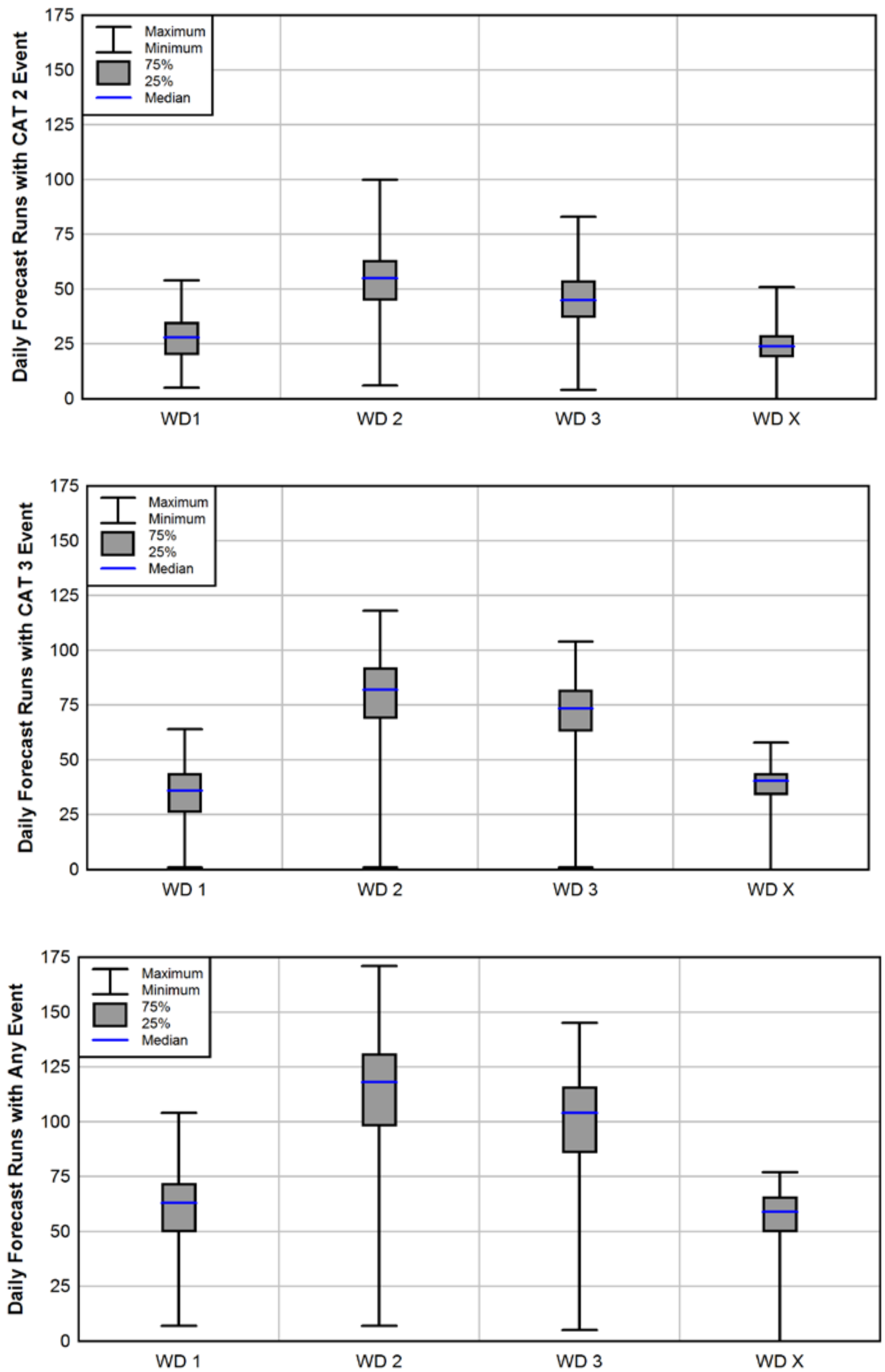


Figure 66. Distributions of the number of daily forecast runs with CAT 2 (top), CAT 3 (middle), or any event (bottom) present in each of the defined Warning Days for the 2011 analysis.

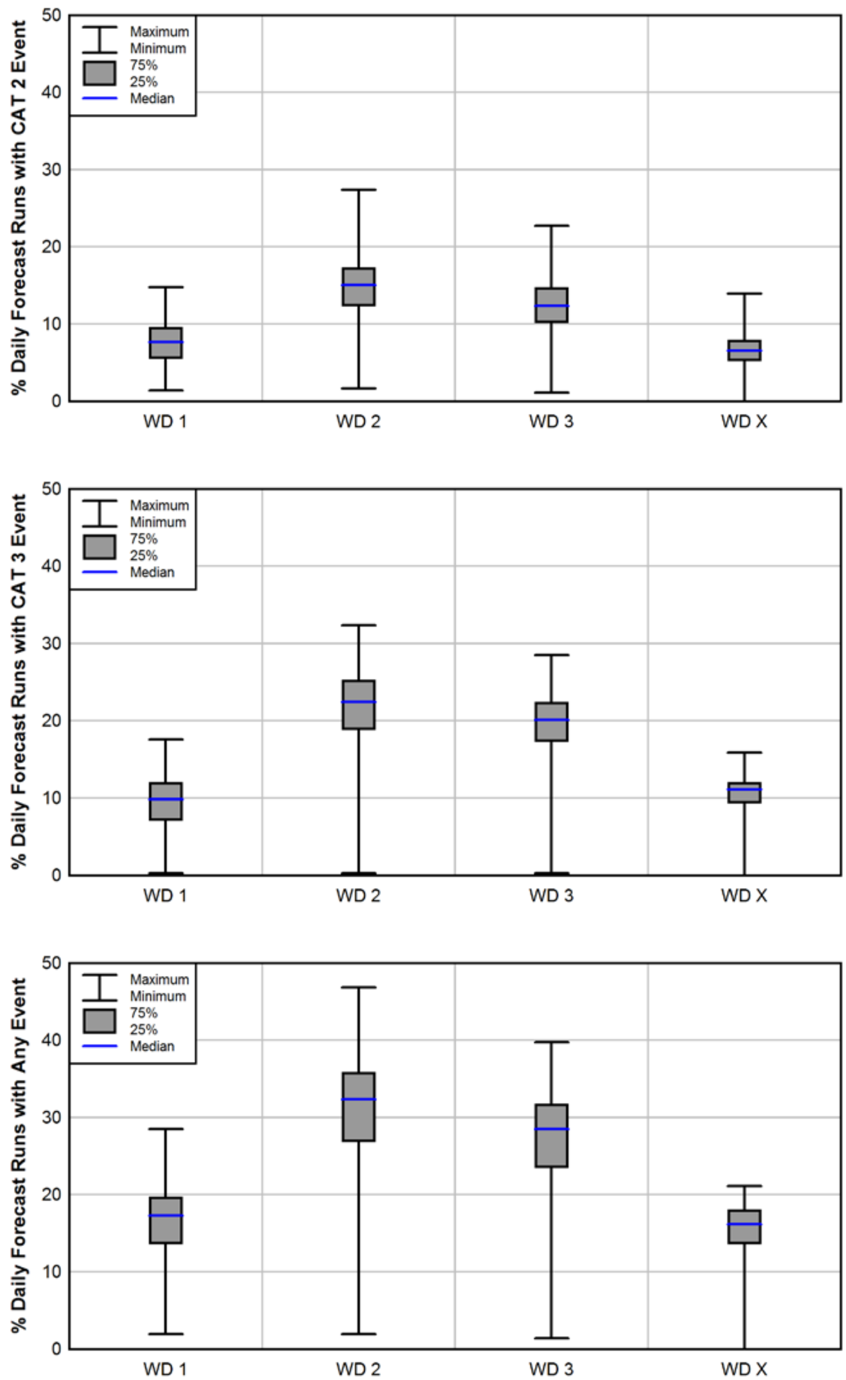


Figure 67. Distributions of the percent of daily forecast runs with CAT 2 (top), CAT 3 (middle), or any event (bottom) present in each of the defined Warning Days for the 2011 analysis.

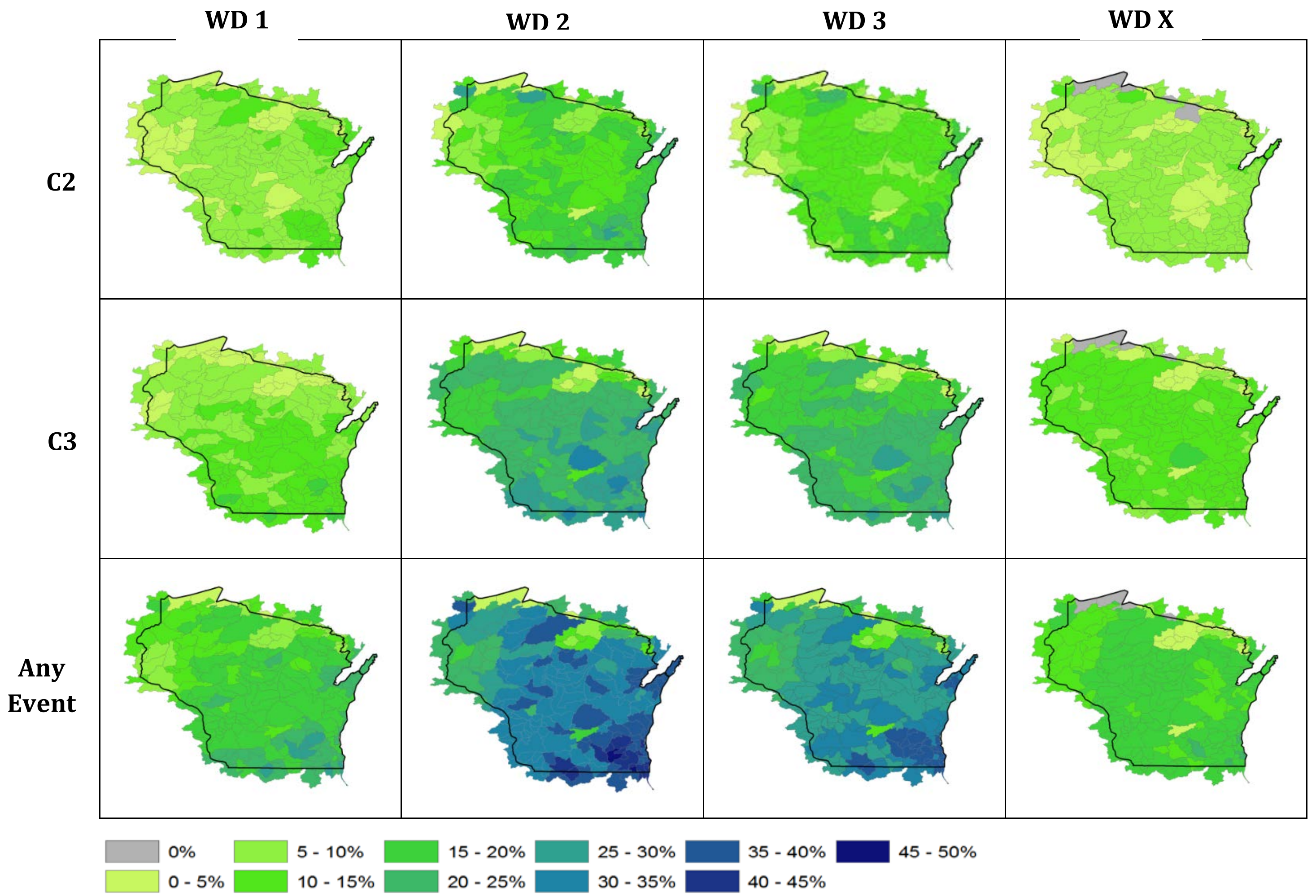


Figure 68. Percent of daily forecast runs with Warning Day runoff event occurrence in the 2011 Forecast analysis. The columns represent the Warning Days periods of 1, 2, 3, and X from left to right. The rows distinguish the event categories with CAT 2, CAT 3, and any event from top to bottom.

4.3.8 Simulated Events Runoff Type Analysis

The last step in the analysis of the 2011 forecast package will examine the breakdown of the runoff events by their specified source of runoff, or runoff type. Recall in section 3.4 it was noted that each runoff event is assigned a value for runoff type indicating if it was caused by rainfall only, a mixture of rainfall and snowmelt, or only snowmelt. In the 2011 forecast analysis 56% of the 40,389 runoff events (any category) in all of the basins were initiated by only rainfall. A mixture was the cause of 19% of the events while 25% were only snowmelt.

On a basin median basis it was found 55% of the events were from rainfall while 20% were from a mix and 25% were from only snowmelt. The ranges for only rainfall events spanned from 18 – 78% whereas a mix of rain and snow spanned 6 – 57% and only snowmelt ranged from 0 – 36% of the events for a given basin. As shown in the top row of Figure 69 below, the median percentage of events by category for a given basin remains dominated by rainfall regardless of CAT 2 or CAT 3. However, the percentages for mix and snowmelt only flip with CAT 3 having nearly four times (28%) more events caused by a mix of rain and snow than those for CAT 2.

The top row of Figure 70 indicates how the median percentages of the events for a given basin breakdown between risk category and runoff type. The figure indicates that CAT 3 rainfall events are in general the most common type of runoff event for a basin. These distributions can be seen spatially in Figure 71 which clearly shows the rainfall only events dominate overall and the southern areas of Wisconsin logically have the highest proportion of only rainfall events. Mix events are much more apparent in CAT 3 than the lower risk category and are more prevalent in central and northern Wisconsin which also passes the logic test. It is also important to note that snowmelt only events are much more prevalent in the CAT 2 category. When high risk snowmelt only events did occur they were found in mostly central Wisconsin.

The second analysis available by runoff type allows examination of the runoff volume by the event sources. The total runoff generated for every basin in the 2011 forecast analysis was 96,975 mm. It was found that 40% was caused by rainfall only, 49% was caused by mixed events, and only 11% was caused by snowmelt only events overall. These same percentages applied to the median basin percentages for runoff volume for all events by runoff types are illustrated by the bottom row of Figure 69. The range of the percentage of total runoff by runoff type spanned 7 – 88% for rainfall only, 8 – 88% for mixed events, and only 0 – 25% for snowmelt only.

As with runoff events, the runoff volume for CAT 2 events on a basin basis is still led by rainfall only with snowmelt only second highest. The mixed events do not occur very often for the lower risk category and therefore do not produce very much of a given basin's CAT 2 runoff (about 8%). Analyzing the high risk events by runoff type identifies the importance of rain on snow events regarding runoff from the landscape. It was shown earlier in this chapter that CAT 3 events make up about 95% of a basin's runoff in the 2011 forecast analysis. Figure 69 below indicates that around half of the CAT 3 runoff is generated from events where rainfall occurs with snowmelt. This finding helps justify the importance state and federal agencies place on situations when these meteorological conditions exist. To further emphasize the importance of mixed runoff events, the bottom row of Figure 70 breaks down the percent of a given basin's runoff by risk category and runoff type. The median value of the highest type and risk is for CAT 3 mixed events with a percentage of almost 47%. The next highest is CAT 3 rainfall only at 36%.

The spatial distribution of event runoff is shown in Figure 72. The focus is primarily on the higher risk category as it overwhelmingly dominates the CAT 2 runoff. Not surprisingly, the basins in far southern Wisconsin are dominated with higher percentage of runoff caused by rainfall only events. However, central through northern Wisconsin is dominated by the mixed source events whereas snowmelt only is a very small percentage. Overall, for the 2011 forecast package, it was found that rainfall only events are the source of the most number of runoff events. However, mixed runoff events dominate the amount of runoff volume.

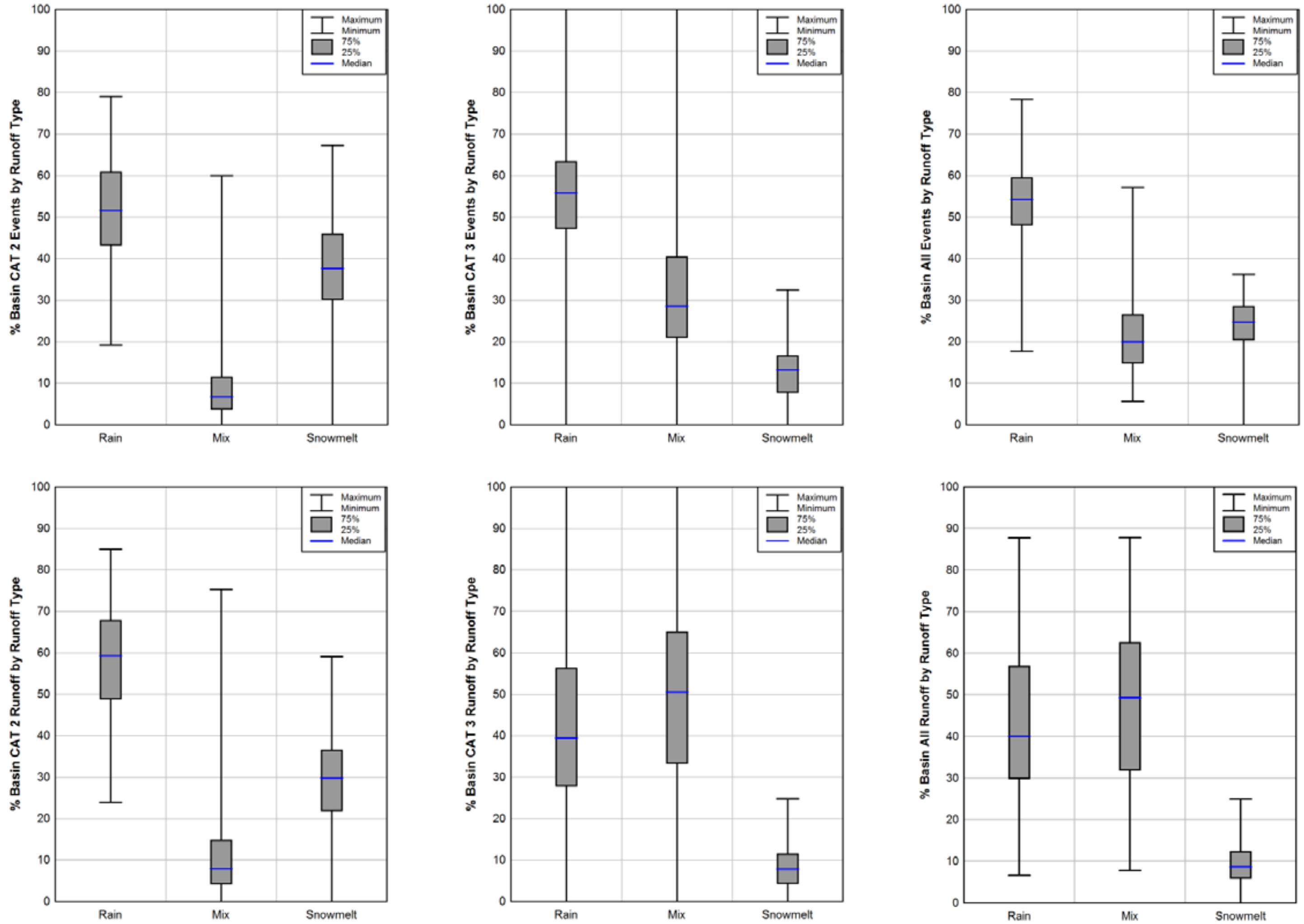


Figure 69. Breakdown of the percent of runoff events (top row) and event runoff (bottom row) by runoff type (only rain, mix, or only snow) for CAT 2, CAT 3, and all events for 216 basins in the 2011 forecast analysis.

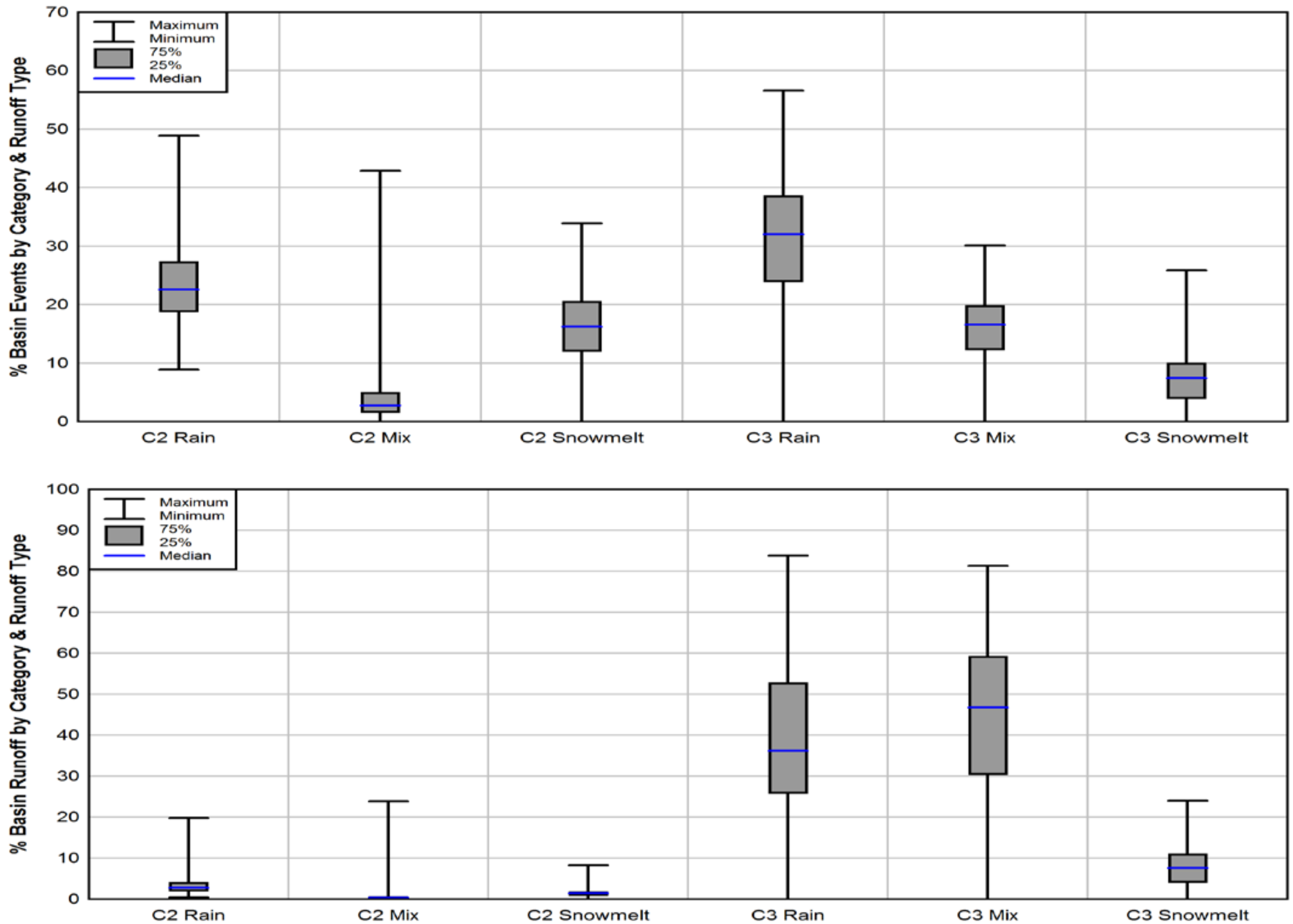


Figure 70. Distribution of the stratification of events (top row) and event runoff (bottom row) for a given basin by risk category and runoff type (only rain, mix, only snowmelt) for the 216 basins in the 2011 forecast analysis.

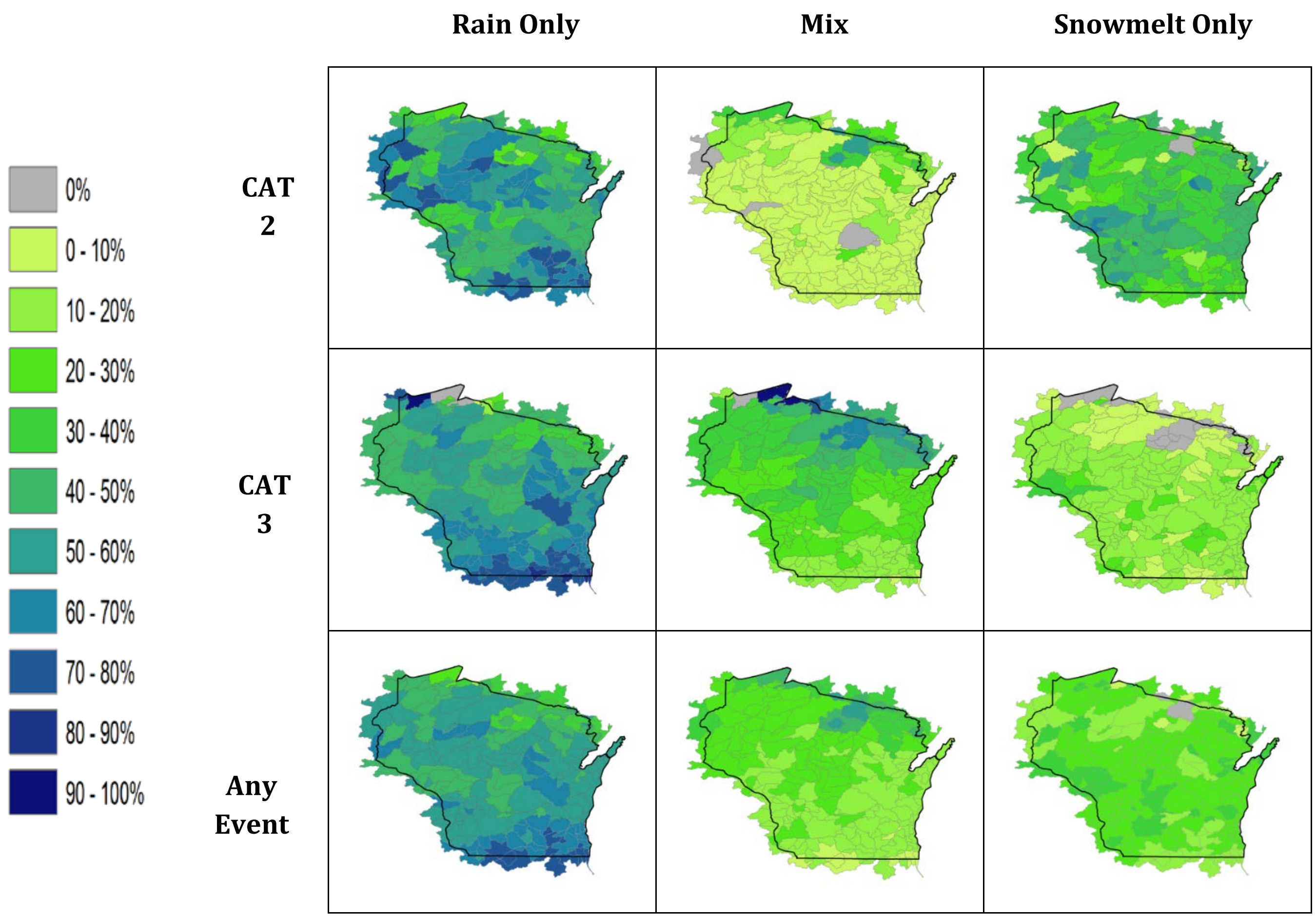


Figure 71. Percentage of runoff events for CAT 2, CAT 3, or any event stratified by runoff event cause (rain only, rain and snowmelt, or only snowmelt) for the 2011 analysis.

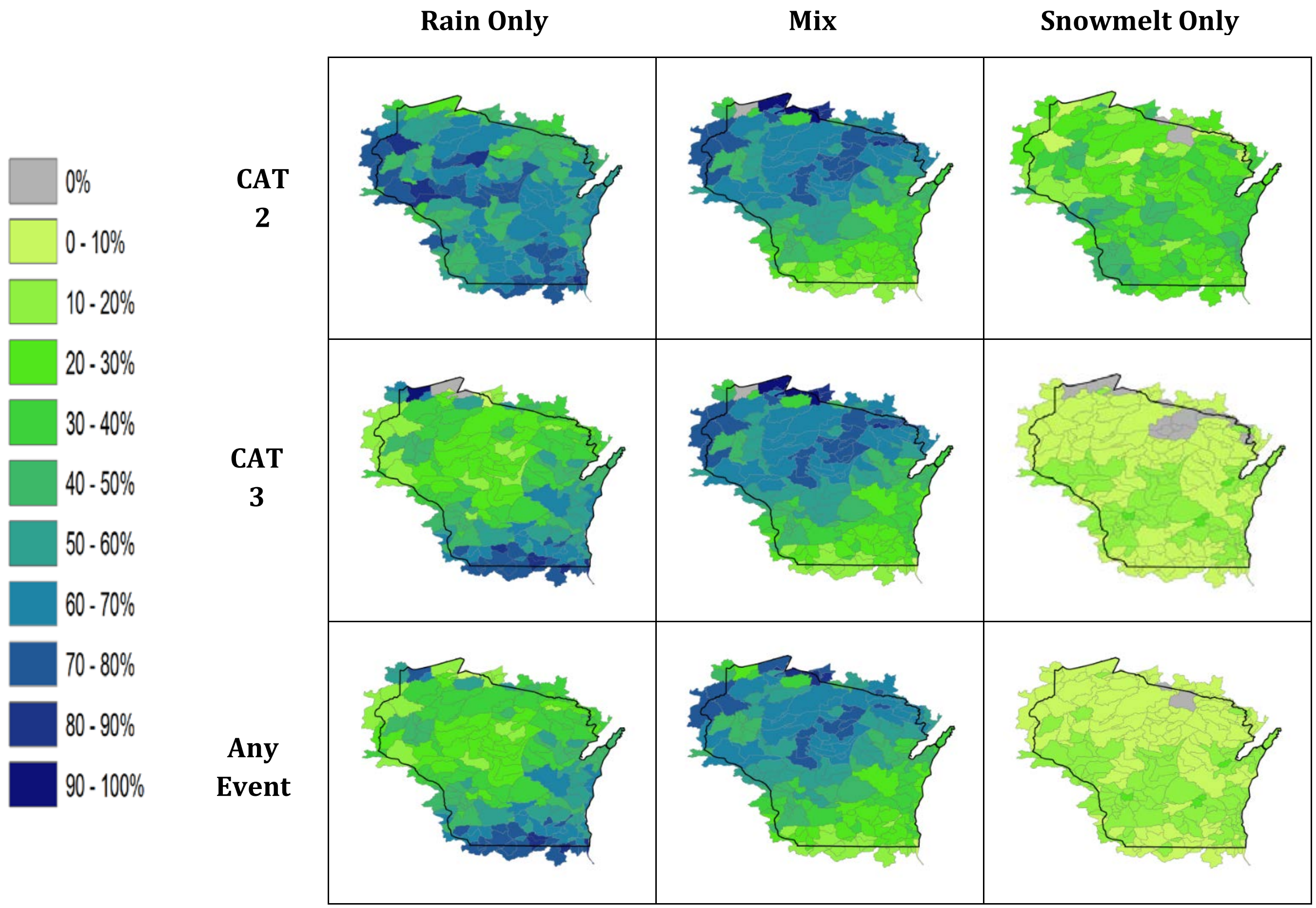


Figure 72. Percentage of event runoff for CAT 2, CAT 3, or any event stratified by runoff event cause (rain only, rain and snowmelt, or only snowmelt) for the 2011 analysis.

4.4 2011 Edge-of-Field Verification

The analysis of RRAF forecast guidance over the 2011 year presented in the preceding sections is important for providing an understanding of the long term behavior and spatial trends across Wisconsin. An equally important summary will be described in this section by highlighting actual verification of the forecast guidance against additional edge-of-field runoff datasets.

Similar to datasets used in Chapter 3 for the RRAF validation, edge-of-field runoff data from Discovery Farms monitored sites across Wisconsin in 2011 were provided by Todd Stuntebeck (USGS). The verification dataset included 9 EOF sites dispersed over 4 NWS basins with each site having an opportunity to record runoff from January 1st through December 31st (Figure 73). The presence of multiple monitoring sites in a given NWS basin introduced the opportunity to combine the runoff events into one time-series for comparison against the forecast guidance for that NWS basin. It should be noted that at this time a dataset of observed runoff events from small USGS gauged watersheds for 2011 was not available for verification analysis.

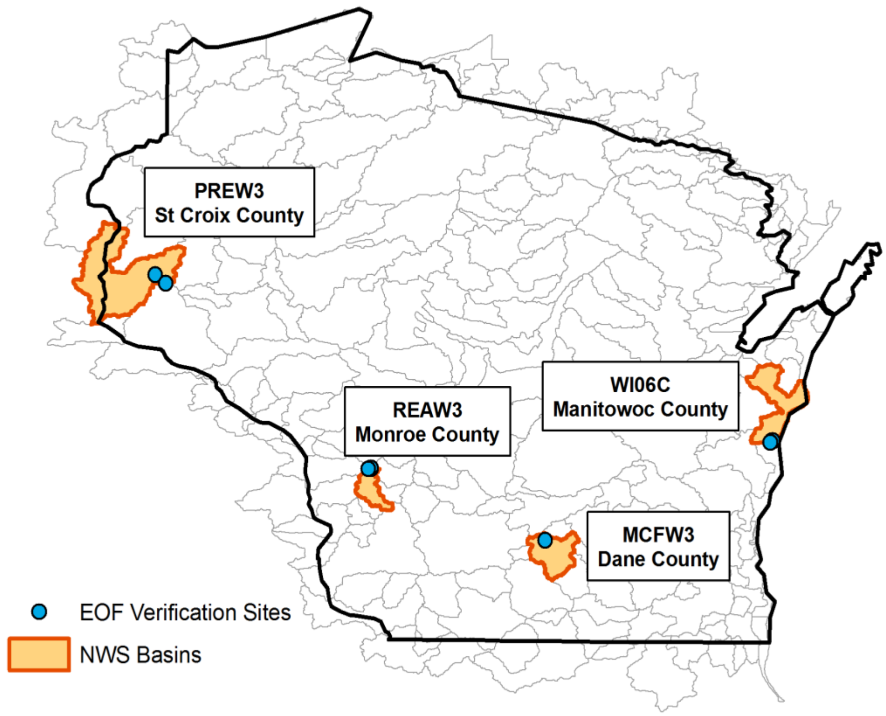


Figure 73. The location of edge-of-field runoff monitoring sites and accompanying NWS basins used in the 2011 RRAF verification.

Areal descriptions of the EOF sites and the corresponding NWS basins are included in Table 23 below. The average EOF site area was found to be 19.7 acres and once again is dwarfed by the average 227,129 acres of the NWS basins. The EOF validation in Chapter 3 included a similar ratio of 21.5 acres for the four EOF sites to 145,468 acres for the NWS basins (Table 2). The average percentage of NWS basin area represented by the individual EOF sites for the validation was found to be 0.015%. For the verification, the ratio was about half as much, near 0.007%, due mostly to the much larger NWS basins involved. However, averaging the ratios of the nine individual EOF sites produces a percent covered value of 0.012%. Regardless of the method, it is clear that only about 0.01% of the NWS basins are covered by the EOF monitoring sites in both the validation and verification.

As mentioned above, the verification process allowed blending up to three EOF sites together for comparison with one host NWS basin. This was done for three counties with multiple EOF sites: Monroe (3 sites), St Croix (2 sites), and Manitowoc (3 sites). This method resulted in an average EOF area of 44.3 acres and when compared to the same NWS average acreage the percentage of NWS basins covered by EOF sites was found to be 0.016%. Averaging the four blended basin ratios produced a value of 0.027%. While either ratio confirms more area of a given NWS basin is represented by the blended EOF approach, the percentages are still extremely small and consistent with the earlier validation.

Table 23. Areal summary of the edge-of-field runoff stations used in the 2011 RRAF verification. The nine EOF stations were dispersed across four NWS basins and averaged about 0.01% of the area of their host NWS basin.

			mi²	km²	acres	% NWS Basin
County EOF	**Site**	**NWS Basins**				
Monroe	5		0.052	0.135	33.4	0.036%
	4	REAW3	0.022	0.057	14.0	0.015%
	2		0.025	0.065	16.0	0.017%
Total			*0.099*	*0.257*	*63.4*	*0.068%*
St Croix	1	PREW3	0.068	0.176	43.4	0.007%
	2		0.013	0.032	8.1	0.001%
Total			*0.081*	*0.208*	*51.5*	*0.009%*
Manitowoc	5		0.024	0.062	15.2	0.006%
	7	WI06C	0.019	0.049	12.0	0.005%
	8		0.008	0.020	5.0	0.002%
Total			*0.051*	*0.131*	*32.2*	*0.013%*
Dane	1	MCFW3	0.047	0.121	30.0	0.017%
EOF Average			***0.031***	***0.080***	***19.7***	***0.012%***
NWS Basins		**EOF Basins**				
REAW3		Monroe	145	375	92,664	
PREW3		St Croix	930	2409	595,277	
WI06C		Manitowoc	384	995	245,870	
MCFW3		Dane	273	707	174,704	
NWS Average			***433***	***1,122***	***277,129***	

4.4.1 Monroe County Edge-of-Field Verification

The first EOF sites to be evaluated are in Monroe County located in southwestern Wisconsin. Three fields were instrumented and used in the verification of the NWS basin REAW3 (92,664 acres). The sites ranged from 14 to 33.4 acres individually while the blended record included the response from all 63.4 acres together. The verification of this EOF-basin pair is summarized in Table 24 and includes statistics for no threshold applied (any event is counted) and threshold applied (only high risk CAT 3 events). Recall the definitions of Bias, CSI, POD are FAR were included in section 3.8 and will be used again for the following verification analysis.

Among the three sites, between 10 and 15 observed runoff events were recorded with the blended record settling on 13 for the year. Interestingly, the two smaller fields in this basin had POD scores of 1.0 whereas the largest field recorded 4 misses out of 15 events for a POD of 0.73. The threshold application had no impact on the perfect POD for the two smaller fields and the largest field picked up one additional miss dropping the POD to 0.67. The blended time-series POD decreased from 0.69 to 0.62 with the threshold inclusion.

Applying the threshold reduced the simulated events by 66 and decreased the FAR by a tenth to 0.63 for the blended record. Even with the threshold in place 63.4% of the simulated events were false alarms (blended) while individually the percentage ranged from 64 – 80%. The observed hit/miss ratios for this verification can be considered questionable due to the small sample of events in the single year evaluated. This is especially true when relying on sample medians as was done in the validation earlier. That stated the hit/miss ratio for observed events with a threshold in place did arrive at 7.5 which continues to show the events hit by the model are considerably larger than the events that were missed. The blended hit/false alarm ratio for the simulated event time-series, before threshold application, was found to be 3.9 and 4.7 for the median and average respectively. Individually the three sites had median ratios between 6.7 and 10.6. The separation described by the ratios continues to support the threshold process as a logical method to remove smaller simulated events that do not correlate with observed response.

As shown by the Bias and CSI scores, the threshold process did improve the accuracy of the REAW3 basin with regards to the three EOF sites. However, a bias score of 6.6 and a CSI score of 0.09 (perfect scores are 1.0 for both) are not strong indicators of a very trustworthy model for this basin. As the POD was fairly good, even perfect at two EOF sites, continued focus on the troubling false alarm behavior in this basin appears to be the main concern.

4.4.2 St Croix County Edge-of-Field Verification

The second verification dataset consisted of two EOF sites located in the headwaters of a very large NWS basin (PREW3) in northwestern Wisconsin. The two EOF sites monitored 8.1 and 43.4 acres for a combined area of 51.5 acres compared to the massive 595,277 acres of the modeled watershed (Table 25). Once again the threshold had no impact on the model's ability to match observed events as site 1 (17 events) had a POD score of 0.59 and site 2 (26 events) scored 0.81 in both scenarios. The blended POD was calculated to be 0.67.

The threshold was a considerable factor in this basin as the number of simulated events dropped 43 to 74 after the application and more importantly the blended FAR decreased 0.18 to a respectable score of 0.27. Therefore, with the threshold, the blended verification indicated 67% of the observed events were hit while only 27% of the simulated events were false alarms. Surprising results considering the immense size differential between the monitored fields and the modeled NWS basin. The post-threshold observed hit/miss ratio was found to be 3.0 and 9.6 for the median and average respectively. The pre-threshold simulated hit/false alarm ratio was determined to be 15.5 and 9.6 for the median and average respectively. Again, both show decent spread between the valued hits and corresponding misses and false alarms.

As alluded to with the mentioned POD and FAR scores, the PREW3 verification had more accurate Bias and CSI scores after the threshold was applied, 1.41 and 0.38 respectively. The performance of the RRAF forecast guidance for PREW3 in 2011 rivals that of the combined scores of the four basins used in the validation (Table 16). An important distinction should be stressed that the validation was completed with events derived from quality-controlled forcings whereas the verification includes forecast precipitation uncertainty by definition making the basin score all that more impressive.

4.4.3 Manitowoc County Edge-of-Field Verification

The selection of Manitowoc County as a verification location is unique in that it was also used during the validation process. However, the three EOF sites evaluated are from a different farm in the same county and NWS basin therefore direct comparison is not entirely valid. The three monitored fields, ranging in size from 5 to 15 acres, recorded between 13 and 22 runoff events in 2011 (Table 26). The NWS model here, an irregularly shaped basin 245,870 acres in size, performed exceptionally well with a blended POD of 0.85 before threshold and 0.81 after. However, individually, POD scores after threshold application ranged from 0.85 to 0.92 which is very good.

The number of simulated events decreased by 84 when the threshold was applied which in turn lowered the blended FAR by 0.14 to a score of 0.45. Therefore with the threshold applied to the blended EOF time-series it was determined that 81% of the observed events were hit by the model and 45% of the simulated events were designated as false alarms. The median observed hit/miss ratio after thresholds were applied was found to be 6.1 while the pre-threshold median simulated hit/false alarm ratio was 7.9.

Applying the threshold to the blended dataset cut the Bias score in half producing a value of 2.50 while the CSI score nearly doubled to 0.30. Both are respectable values yet room for improvement hinges on reducing false alarms as the POD is already a decent score for this basin.

4.4.4 Dane County Edge-of-Field Verification

The final RRAF verification completed involved one EOF site located in south-central Wisconsin in Dane County. The 30 acre site, located in the 174,704 acre MCFW3 basin, recorded 12 observed hits in 2011 with 11 of them being hits without a threshold and only 10 with the threshold (Table 27). This breakdown provided a POD of 0.92 and 0.83 before and after the threshold.

This MCFW3 models are fairly responsive with regards to the RRAF as 228 events were simulated initially. After threshold application that number dropped to 136 indicating many were very small in magnitude. The high number of simulated events led to a high FAR of 0.82 and 0.76 before and after threshold usage. It seems clear that this basin captures real events when they occur but also produces a very high number of small events that are considered false alarms.

Due to the small sample size the median observed hit/miss ratios are not useful. However, the average ratios are 24 and 18 before and after thresholds. The median simulated hit/false alarm ratio was a respectable 3.8 before thresholds. Even though this basin does a great job capturing observed runoff occurrences, the extremely high number of false alarms dampens its accuracy and therefore its effectiveness. A very troubling Bias score of 9.50 and CSI of 0.09 after threshold application represent that fact.

Table 24. Verification of 2011 RRAF guidance for the REAW3 basin with EOF runoff sites in Monroe County. A summary is provided with all events taken into account as well as with just CAT 3 events above the basin threshold.

Monroe County	No Threshold Applied				Threshold Applied			
REAW3	**Site 2**	**Site 4**	**Site 5**	**Blended**	**Site 2**	**Site 4**	**Site 5**	**Blended**
Acres	16.0	14.0	33.4	**	16.0	14.0	33.4	**
Obs Events	10	8	15	13	10	8	15	13
Obs Hits	10	8	11	9	10	8	10	8
Obs Miss	0	0	4	4	0	0	5	5
Sim Events	189	189	189	189	123	123	123	123
Sim Hits	27	32	48	51	25	32	44	45
Sim FA	162	157	141	138	98	91	79	78
Bias	17.2	20.6	10.1	11.3	10.8	12.4	5.9	6.6
POD	1.00	1.00	0.73	0.69	1.00	1.00	0.67	0.62
FAR	0.86	0.83	0.75	0.73	0.80	0.74	0.64	0.63
CSI	0.06	0.05	0.07	0.06	0.09	0.08	0.11	0.09
% Obs Evts = Hit	100%	100%	73.3%	69.2%	100%	100%	66.7%	61.5%
% Obs Evts = Miss	0%	0%	26.7%	30.8%	0%	0%	33.3%	38.5%
% Sim Evts = Hit	14.3%	16.9%	25.4%	27.0%	20.3%	26.0%	35.8%	36.6%
% Sim Evts = FA	85.7%	83.1%	74.6%	73.0%	79.7%	74.0%	64.2%	63.4%
Median Obs Hit/Miss	**	**	0.3	0.3	**	**	1.2	7.5
Avg Obs Hit/Miss	**	**	2.5	5.3	**	**	3.4	7.4
Median Sim Hit/FA	6.7	10.6	8.8	3.9	3.5	3.5	3.2	1.5
Avg Sim Hit/FA	6.8	6.1	4.9	4.7	4.7	3.7	3.2	3.2

Table 25. Verification of 2011 RRAF guidance for the PREW3 basin with EOF runoff sites in St Croix County. A summary is provided with all events taken into account as well as with just CAT 3 events above the basin threshold.

St Croix County	No Threshold Applied			Threshold Applied		
PREW3	**Site 1**	**Site 2**	**Blended**	**Site 1**	**Site 2**	**Blended**
Acres	43.4	8.1	**	43.4	8.1	**
Obs Events	17	26	27	17	26	27
Obs Hits	10	21	18	10	21	18
Obs Miss	7	5	9	7	5	9
Sim Events	117	117	117	74	74	74
Sim Hits	49	64	64	44	54	54
Sim FA	68	53	53	30	20	20
Bias	4.59	2.85	2.63	2.35	1.58	1.41
POD	0.59	0.81	0.67	0.59	0.81	0.67
FAR	0.58	0.45	0.45	0.41	0.27	0.27
CSI	0.12	0.27	0.23	0.21	0.46	0.38
% Obs Evts = Hit	58.8%	80.8%	66.7%	58.8%	80.8%	66.7%
% Obs Evts = Miss	41.2%	19.2%	33.3%	41.2%	19.2%	33.3%
% Sim Evts = Hit	41.9%	54.7%	54.7%	59.5%	73.0%	73.0%
% Sim Evts = FA	58.1%	45.3%	45.3%	40.5%	27.0%	27.0%
Median Obs Hit/Miss	78.5	1.7	3.0	78.5	1.7	3.0
Avg Obs Hit/Miss	31.4	4.0	9.6	31.4	4.0	9.6
Median Sim Hit/FA	11.9	15.5	15.5	8.7	5.9	5.9
Avg Sim Hit/FA	9.1	9.6	9.6	5.2	5.5	5.5

Table 26. Verification of 2011 RRAF guidance for the WI06C basin with EOF runoff sites in Manitowoc County. A summary is provided with all events taken into account as well as with just CAT 3 events above the basin threshold.

Manitowoc County	No Threshold Applied				Threshold Applied			
WI06C	**Site 5**	**Site 7**	**Site 8**	**Blended**	**Site 5**	**Site 7**	**Site 8**	**Blended**
Acres	15.2	12.0	5.0	**	15.2	12.0	5.0	**
Obs Events	22	13	20	26	22	13	20	26
Obs Hits	19	12	18	22	19	12	17	21
Obs Miss	3	1	2	4	3	1	3	5
Sim Events	182	182	182	182	98	98	98	98
Sim Hits	68	51	61	74	50	40	46	54
Sim FA	114	131	121	108	48	58	52	44
Bias	6.1	11.0	7.0	5.0	3.1	5.4	3.5	2.5
POD	0.86	0.92	0.90	0.85	0.86	0.92	0.85	0.81
FAR	0.63	0.72	0.66	0.59	0.49	0.59	0.53	0.45
CSI	0.14	0.08	0.13	0.16	0.27	0.17	0.24	0.30
% Obs Evts = Hit	86.4%	92.3%	90.0%	84.6%	86.4%	92.3%	85.0%	80.8%
% Obs Evts = Miss	13.6%	7.7%	10.0%	15.4%	13.6%	7.7%	15.0%	19.2%
% Sim Evts = Hit	37.4%	28.0%	33.5%	40.7%	51.0%	40.8%	46.9%	55.1%
% Sim Evts = FA	62.6%	72.0%	66.5%	59.3%	49.0%	59.2%	53.1%	44.9%
Median Obs Hit/Miss	1.3	0.3	1.0	3.9	1.3	0.3	4.8	6.1
Avg Obs Hit/Miss	1.2	0.5	0.9	1.4	1.2	0.5	1.4	1.8
Median Sim Hit/FA	7.9	8.1	8.4	7.9	1.9	2.1	2.1	2.1
Avg Sim Hit/FA	2.2	2.3	2.4	2.5	1.3	1.3	1.4	1.5

Table 27. Verification of 2011 RRAF guidance for the MCFW3 basin with EOF runoff sites in Dane County. A summary is provided with all events taken into account as well as with just CAT 3 events above the basin threshold.

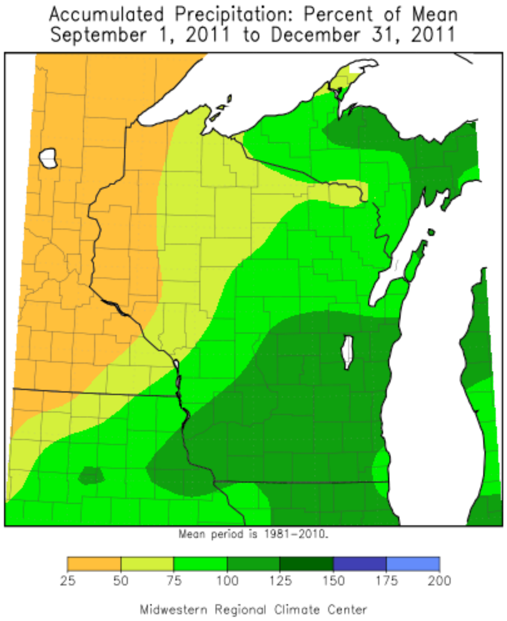
Dane County	No Threshold Applied	Threshold Applied
WI06C	Site 1	Site 1
Acres	30.0	30.0
Obs Events	12	12
Obs Hits	11	10
Obs Miss	1	2
Sim Events	228	136
Sim Hits	40	32
Sim FA	188	104
Bias	16.58	9.50
POD	0.92	0.83
FAR	0.82	0.76
CSI	0.06	0.09
% Obs Evts = Hit	91.7%	83.3%
% Obs Evts = Miss	8.3%	16.7%
% Sim Evts = Hit	17.5%	23.5%
% Sim Evts = FA	82.5%	76.5%
Median Obs Hit/Miss	0.2	0.3
Avg Obs Hit/Miss	24.4	17.7
Median Sim Hit/FA	3.8	2.0
Avg Sim Hit/FA	2.5	1.8

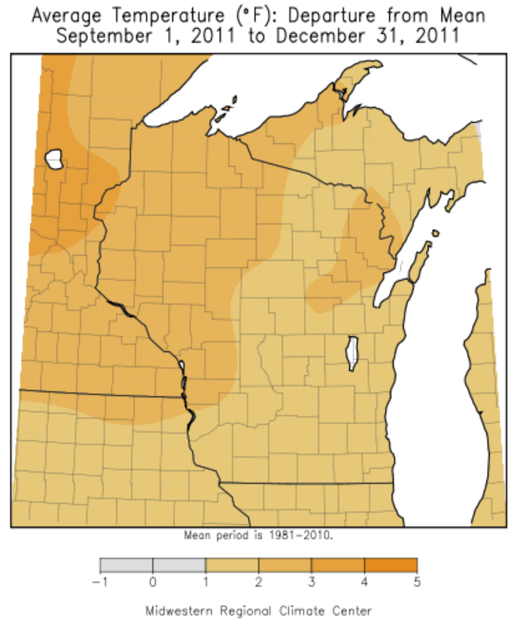
4.4.5 2011 Edge-of-Field Verification Summary

Wanting to generalize the RRAF performance during the 2011 forecast season the median or average summary statistics were chosen from the nine EOF sites and 4 NWS basins. Two summary methods were completed. The first treated the 9 verification studies individually and chose statistics from that sample (Table 28). The second method combined the Dane County data with the blended values for Monroe, St Croix, and Manitowoc Counties to test the possibility that blended values from several EOF sites for a given NWS basin would be more representative of the RRAF performance for that basin (Table 29). As stated earlier, the validity of this verification is stretched by several factors: (1) the short time frame (1-year), (2) the small sample of observed events at any given EOF site (range of 8 – 26), and (3) the fact that the 2011 fall and winter in Wisconsin were warmer with much below normal snowfall across the southern two-thirds of the state (Figure 74). It is interesting to note that none of the nine EOF sites recorded any runoff events after mid-September.

Table 28. Summary of the 2011 RRAF verification using the values of nine EOF sites across four NWS basins.

All Sites (Individual)	**No Thresh**	**Thresh**	**T - NT**	**% Change**
Median Obs Hit/Miss Ratio	1.31	1.71	0.40	30.1%
Average Obs Hit/Miss Ratio	9.27	8.52	-0.75	-8.1%
Median Sim Hit/FA Ratio	8.36	3.18	-5.18	-62.0%
Average Sim Hit/FA Ratio	5.11	3.13	-1.97	-38.7%
Median Bias	10.13	5.38	-4.75	-46.9%
Median CSI	0.08	0.17	0.09	102.8%
Median POD	0.90	0.85	-0.05	-5.6%
Median FAR	0.72	0.59	-0.13	-17.8%
Median % Obs Hit	90.0%	85.0%	-5.0%	-5.6%
Median % Obs Miss	10.0%	15.0%	5.0%	50.0%
Median % Sim Hit	28.0%	40.8%	12.8%	45.7%
Median % Sim FA	72.0%	59.2%	-12.8%	-17.8%





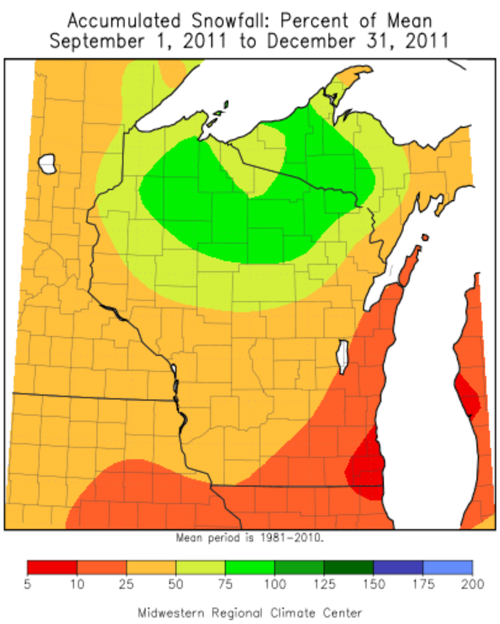


Figure 74. 2011 Wisconsin climate summaries for September through December. Upper Left image is percent of mean precipitation for that time period. Upper right image is departure from mean temperature. Lower left image is percent of mean accumulated snowfall for this time frame. Images courtesy of Midwestern Regional Climate Center.

The median observed hit/miss ratio for the non-blended method was a weak 1.7 value after thresholds whereas the average ratio was a more reasonable 8.5. The median POD across all EOF sites was a great 0.90 initially and a very respectable 0.85 after thresholds were applied. The pre-threshold FAR measured 0.72 and only decreased to 0.59 after thresholds were introduced. This score indicates that 60% of simulated events were considered false alarms in these basins validating initial concern in the project and the motivation for thresholds to be introduced into the process. The pre-threshold simulated hit/false alarm ratio of 8.36 suggests a strong separation between the important hit events and the false alarms existed.

As witnessed for the individual analyses, the Bias and CSI score do improve for the non-blended median of the verification sample once thresholds are used. However, the scores (5.4 for Bias and 0.17 for CSI) do suggest some work is needed to improve the accuracy and reliability of the RRAF by reducing the drag from too many false alarms.

Table 29. Summary of the 2011 RRAF verification for four NWS basins using blended EOF values for counties with more than one EOF site.

Median All Sites (Blended)	No Thresh	Thresh	T - NT	% Change
Median Obs Hit/Miss Ratio	3.43	6.29	2.86	83.6%
Average Obs Hit/Miss Ratio	10.18	9.14	-1.04	-10.2%
Median Sim Hit/FA Ratio	5.90	2.06	-3.84	-65.1%
Average Sim Hit/FA Ratio	4.84	3.00	-1.84	-37.9%
Median Bias	8.15	4.56	-3.60	-44.1%
Median CSI	0.11	0.19	0.08	73.3%
Median POD	0.77	0.74	-0.03	-4.2%
Median FAR	0.66	0.54	-0.12	-18.2%
Median % Obs Hit	76.9%	73.7%	-3.2%	-4.2%
Median % Obs Miss	23.1%	26.3%	3.2%	13.9%
Median % Sim Hit	33.8%	45.8%	12.0%	35.5%
Median % Sim FA	66.2%	54.2%	-12.0%	-18.2%

Recall that the second method summary incorporated the blended EOF verification values from three of the four county test locations. The observed hit/miss and simulated hit/false alarm ratios were all reasonable and continue to indicate stratification between the desired hits and undesirable misses and false alarms. The POD values for this method came in a little lower compared to the non-blended approach with 0.77 and 0.74 for before and after thresholds. On the other hand, the FAR scores were a bit better in both scenarios at 0.66 and 0.54 for before and after thresholds. The composite scores were also a bit better with the blended approach as the median Bias was 4.56, almost down a point. The CSI after thresholds was 0.19, up slightly compared to the first method. Overall, the blended median approach does produce a slightly better evaluation of the RRAF verification, due again to the lower false alarm rate.

4.5 RRAF Performance Summary & Adjustments

Up to this point the Runoff Risk Advisory Forecast decision support tool has been described and evaluated in many ways. This section will provide an overall summary of performance as well as introduce the latest modifications that will be applied to further enhance the product accuracy.

4.5.1 Detection of Observed Events by the RRAF

Throughout all the RRAF evaluations completed in this study it was determined the model was very impressive at detecting observed runoff events when they occur. Detection rates for the EOF validation were found to be 80% while the USGS comparison was 62%. The EOF verification was very similar with a detection rate of 77% (Table 30). When basin thresholds were applied to create the high risk category the probability of detection did decrease between 16% – 21% for the validation comparison and only 3% for the EOF verification. However, the argument is presented that since the events missed due to the threshold application are much smaller than the events hit, the threshold process actually results in a better RRAF product despite the lower POD values.

An important aspect of the impressive detection rates that needs to be remembered is that the observed data for the EOF comparisons represented only 0.01% of the modeled basin areas while the USGS comparison was nearly 5%. It was also reassuring to find that the model capability for capturing observed events was found not only in the validation of historical simulated data but also in the verification of daily forecast runs where uncertainty in precipitation and temperatures was introduced.

Regarding the high risk category, the lower POD scores could be a cause of concern in terms of RRAF accuracy. Missed events demanded a lot of attention from the Working Group during model evaluation as the general consensus is to of error on the side of caution to avoid potential contamination incidents with this decision support tool. With that said, the following supports the argument that thresholds help produce a better forecast tool for the producers even if the POD is slightly lower for the higher risk events.

Table 30. Summary of median POD and FAR scores for the EOF and USGS validation and the blended EOF verification.

RRAF Summary of Median POD and FAR Scores for Validation and Verification Analyses				
	Probability of Detection (POD)		**False Alarm Rate (FAR)**	
	No Thresh	**Thresh**	**No Thresh**	**Thresh**
Validation				
EOF	0.80	0.64	0.71	0.48
USGS	0.62	0.41	0.45	0.19
Combined	0.71	0.53	0.58	0.34
Verification				
Blended EOF	0.77	0.74	0.66	0.54

Table 31. Summary of the median observed runoff event hit/miss ratio and simulated runoff event hit/false alarm ratio for the EOF and USGS validation and the blended EOF verification. Values in parentheses are the blended EOF average values for comparison due to small sample size effect on median verification ratios.

RRAF Summary of Observed and Simulated Runoff Event Ratios for Validation and Verification Analyses				
	Observed Hit/Miss Ratio		**Simulated Hit/False Alarm Ratio**	
	No Thresh	**Thresh**	**No Thresh**	**Thresh**
Validation				
EOF	5.0	4.9	10.3	1.4
USGS	10.5	14.2	6.2	1.3
Combined	7.7	9.6	8.3	1.4
Verification				
Blended EOF	3.4 (10.2)	6.3 (9.1)	5.9	2.1

The key argument for this assumption involves highlighting the observed hit/miss ratios in Table 31. The table shows that the magnitude of events captured by the model ranged from 5 – 10 times larger than the missed observed events for the validation. Applying thresholds resulted in either the same or even higher ratios. A similar pattern was seen in the verification comparison. Since the POD decreases with threshold application while the ratios either remain steady or increase, it can be inferred that the thresholds are only eliminating simulated events that corresponded with the smallest observed events.

Recall in Chapter 2, it is often found that most of the annual nutrient transport from fields occurs during a few larger events. The hit/miss ratios suggest the high risk category is not negatively affected by removing smaller simulated events. The high risk category is still capturing the largest events of greatest concern, and therefore, the POD decrease for CAT 3 is acceptable. The above discussion elaborates on why the thresholds are useful for focusing producers on the most important runoff events. The second and equally important consequence of thresholds is to lower false alarms. This is the topic of the next sub-section.

4.5.2 The Challenge of False Alarms

Early on in the RRAF development it was found that false alarms could be a significant challenge due to their elevated occurrence rate. Developing and implementing the basin specific thresholds helped focus the largest events into the high risk category. By ignoring the smallest 60% of simulated events, this reduced false alarms in the high risk category and will help build credibility in the RRAF.

As seen in Table 30, the initial EOF validation tallied a 71% false alarm rate. The USGS comparison did produce a lower occurrence with only 45%. The resultant CAT 3 false alarm rates were found to be 48% and 19% for the EOF and USGS validations respectively which were decent reductions of 23% and 26%. The EOF verification presented an initial false alarm rate of 66% which was reduced by only 12% after thresholds were used and only high risk was evaluated. Note the future precipitation (QPF) influence could be affecting the lower drop in the EOF verification.

At first glance of the CAT 3 false alarms (48%, 19%, 54%, for EOF validation, USGS validation, and EOF verification respectively), one could justify an approach where the basin thresholds are optimized to maximize the summary statistics. However that was not the approach chosen in this study or by the RRAF Working Group. The challenge of high false alarms must be kept in perspective so that informed decisions can be made to produce the best overall decision support tool. To keep this perspective two factors must be balanced: (1) the desire for acceptable results and (2) the limitations of the model comparisons.

The high risk category does improve the model results even though the resultant false alarms are still a bit high for the EOF comparisons. However, it was decided no changes to the threshold process would occur mainly due to one important limitation inherent in the comparison analyses: the immense spatial scale difference between the observed datasets and the modeled basins. Several examples are provided below that revolve around doubt this issue can interject into the analysis which could artificially enhance the false alarm rates.

The first point involves clearly understanding what is not being sampled during the RRAF validations and verification. There is an immense area of the NWS basins (95 – 99.99%) not being accounted for by the observed datasets. One could safely propose that some false alarms indicated by the RRAF might be corresponding to actual runoff occurring in parts of the modeled basin that are not monitored. Further, recall in the EOF verification where multiple EOF sites were located within short distances on the same farm. There was variability between the sites and blending the sites together generally produced a better comparison with the model. This point also supports caution with making decisions based on only one observation plot in an entire NWS basin.

The USGS comparison included more area than the EOF comparisons and recorded much lower false alarm rates. This was theorized as more land area and variability was averaged into these small watersheds. In addition, the USGS comparison still only covered 5% of the NWS basins. It could also be proposed that the trend of decreasing false alarm rates would continue if a higher percentage of the NWS basins were overlapped by observed data.

The possible doubt in how much of the actual observed response is captured in the RRAF comparisons combined with the desire to error on warning too often instead of not enough led the Working Group to remain steady with the high risk basin thresholds. It appears there is an acceptable balance between the summary statistics, including the false alarm rate, and the known limitations illustrated above.

One additional point can be proposed that supports the Working Group's stance of not wanting to modify basin thresholds or totally downplaying the CAT 2 (moderate risk) events. This argument is not something explicitly taken into account by the models behind the RRAF, however does incorporate actual manure application practices. The concern involves unaccounted water that is applied to fields as liquid dairy manure (LDM) slurry. Discovery Farms states that 7,000 (13,500) gallons of LDM applied per acre is equivalent to a 0.25 (0.5) inch rainfall event (UW-Extension, 2011 [a]) (Discovery Farms, 2011).

As mentioned, this water content is not included in the RRAF modeling. However, runoff risk can still be applicable to a producer who knows they will be applying LDM to their fields. If the RRAF is indicating moderate risk in that basin, the models suggest conditions are favorable for runoff to occur. Perhaps runoff does not occur naturally in that basin from rainfall or snowmelt and therefore is considered a false alarm. However, if a farmer adds 0.25 – 0.5” of water to that field the RRAF risk suggests he probably will overwhelm that field and runoff will occur. This is another example of how that producer can use the RRAF as a decision support tool to avoid applying LDM even during marginally risky conditions.

4.5.3 Latest RRAF Modification

The previous sub-sections discussed the RRAF ability to replicate observed runoff events as well as the benefits of the threshold technique applied. Members of the Working Group feel comfortable with the process of the RRAF creating two risk categories where the most emphasis is placed on the high risk CAT 3 events. It is also firmly believed this first-generation tool can be continually evaluated and improved. The most recent modification that will be implemented in the spring of 2013 is aimed at reducing false alarms caused by the smallest simulated events.

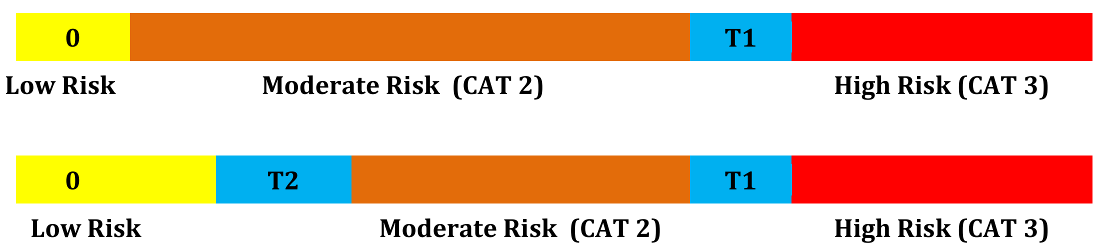
The new modification will introduce a second, lower threshold for each basin which will create a new category of simulated events. This new class of events will be essentially ignored and removed from the CAT 2 (moderate risk) category (Table 32). As mentioned above, the Working Group is comfortable with the justification behind the original threshold and the high risk events. Further, there is comfort in still promoting most of the moderate risk events due to the comparison limitations described above. However, this new process is an attempt at optimizing the product by removing some of the smallest simulated events which have been found to be mostly false alarms. Essentially this action will tighten up the runoff risk focus and concentrate producers on the conditions most likely suitable for runoff to occur in their area. It will be shown that these new lower thresholds remove the smallest simulated events producing minimal impact on the validation and verification of the RRAF model. In addition, credibility can be built in the tool by eliminating the false alarms associated with these small events.

A similar method to the one used earlier in section 3.11 for determining the original basin specific thresholds was followed for the second threshold application. The process started by comparing the value of the smallest simulated hit for each NWS basin to the distribution of simulated events derived from each basin's 50-year historical simulation. The goal was to evaluate for a pattern in the exceedance of that runoff event magnitude for each of the basins.

The USGS watershed validation revealed no clear signal as extremely small events were also being typed as hits in this comparison. However, the EOF validation suggested a pattern could be found. The exceedance probabilities of the smallest simulated hits for the four NWS basins ranged from 0.72 – 0.88 with average 0.799. Therefore an exceedance probability of 0.80 was chosen as the new lower threshold and a list of corresponding simulated event magnitudes was generated for the basins. With regards to the historical simulated event distribution, the smallest 20% of events would now be ignored and fall into the low risk category (no event). The middle 40% (between 0.80 and 0.40 exceedance) will be categorized as moderate risk, and the top 40% remains the high risk category.

Table 32. Schematic representation of the category definitions of runoff risk for the current RRAF (top row) and the change proposed for spring 2013 (bottom row) where a second lower threshold (T2) will be introduced. The x-axis represents simulated runoff event magnitude starting at zero or no event on the left.

Representation of the Current and Proposed Runoff Risk Categories



The next step in the evaluation of the lower threshold was to re-run the validation and verification comparisons with the lowest events removed. In the EOF validation of the four NWS basins an average of 25 false alarms were removed for nearly a 20% decrease. A 1% decrease in the number of simulated hits was witnessed. Overall the average POD for the basins decreased 0.6% while the FAR decreased 12.6% so a definite improvement with little cost was obtained. In the USGS validation an average of 59 false alarms were removed (28% decrease) while 34 hits were removed (13% decrease). The overall average POD decreased 6.8% while the average FAR dropped 12.1%. Although not as clean of an improvement, a benefit is still found.

For the EOF verification of the 2011 forecast runs minimal impact was observed by the new lower threshold. The average POD did not change while the average FAR only saw a 2.3% decrease. It is likely premature to make a final judgment on the second threshold effectiveness with regards to the forecast verification due to the small sample size associated with the single year studied.

4.6 NCRFC Model Adjustment for RRAF Performance

The 2011 analysis of the daily forecast runs helped identify several basins that produced inconsistent RRAF guidance compared to neighboring basins in their region. Specifically, these basins were not generating enough runoff events when compared to nearby basins. Since the quantity of runoff events (regardless of risk category) is not a function of the threshold process, a more in-depth evaluation of the basin model parameters was required.

Several of the problem basins were located along the northern shore of Wisconsin and drain directly into Lake Superior. Having no single defined outlet and encompassing several small streams, these basins were provided estimated model parameters when they were initially defined in the NCRFC area. Operationally, these basins have minimal impact on streamflow forecasting as they do not route water to downstream basins. The RRAF evaluation determined that these basins could be re-calibrated and given new regionalized parameters based on more trustworthy basins nearby. This was done for the following basins: WI12C, WI13C, WI14C, and WI15C. Before this change these four basins generated very few runoff events.

Probably the basin that stuck out the most in terms of RRAF spatial consistency was PDSW3 down in south central Wisconsin. The parameters for this basin built up a much larger UZTW deficit as well as smaller amounts of interflow runoff over the 2011 forecast analysis. This resulted in fewer runoff events and often times this basin would be indicating no risk while every basin around it simulated runoff events. Located along the Wisconsin River near Prairie du Sac, this basin is dominated by a reservoir at its outlet. The local runoff is generally much smaller in magnitude compared to the routed flow to this location therefore local runoff calibration was difficult. Again, this basin was re-evaluated and new basin parameters were derived based on nearby basins. This basin now behaves consistently with its region.

Finally, the basin MRNM4 straddling the Michigan-Wisconsin border is another example that provides very few runoff events compared to nearby basins. In fact it produces almost 99% less interflow when compared to the neighbor average. The model parameters were evaluated and unfortunately it was determined it cannot be changed at this time. This basin has a dam upstream of the outlet which is not modeled explicitly. To account for the delay in local runoff being routed to downstream basins the SAC model was initially calibrated to force most of the runoff to be distributed as base flow. Therefore this basin does not generate much interflow at all and thus few events are simulated for the RRAF. As this basin is not in a heavy dairy or agriculture area, it seems allowing it to continue as an outlier has minimal impact at this time.

For the basins that were re-calibrated new historical simulations will be required. The new historical simulations will be used to generate a list of runoff events for that duration and allow new basin thresholds to be determined. The five basins re-calibrated for the RRAF product are shown in Figure 75 below.

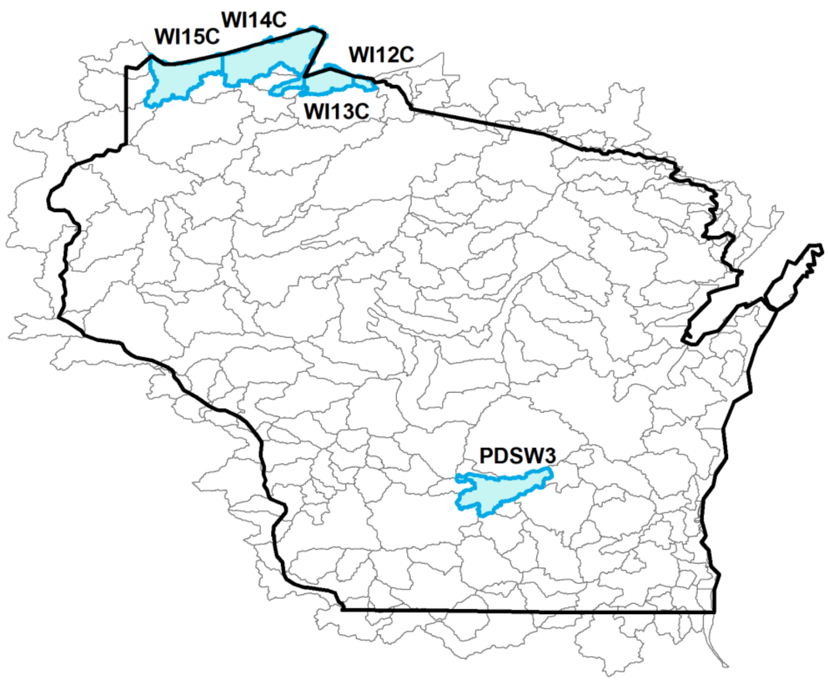


Figure 75. The five NWS basins (shaded in blue) that were re-calibrated to allow for spatially consistent RRAF response.

At this time no other modifications have been made to the NWS models or RRAF definition. The Working Group will continue to evaluate the guidance and it is possible alterations to the threshold process could be changed overall or on an individual basin basis.

4.7 Future Steps for RRAF Guidance

Over the last few years the RRAF Working Group has periodic meetings to discuss the status, performance, web interface, and other aspects of the decision support tool. With the ultimate goal to provide the best decision support guidance to manure and fertilizer applicators in Wisconsin, the team routinely self-evaluates the RRAF webpage and forecasts with the desire to improve accuracy, reliability, and clarity for the end user. Members of the group have also increased efforts to obtain user feedback whenever the RRAF is promoted in outreach events and media.

In terms of outreach, the NCRFC has been proactive in briefing fellow agencies and educational institutions about the RRAF concept. Overall, feedback from the varying audiences has been very positive. Some of the federal and state agencies and organizations briefed by the NCRFC to this point include:

- U.S. Geological Survey, Midwest Region
- U.S. Army Corps of Engineers, Mississippi Valley Division
- Environmental Protection agency, Region V
- Natural Resource Conservation Service, Midwest Region
- NOAA/National Ocean Service – Center for Sponsored Coastal Ocean Research (Gulf of Mexico Hypoxia Task Force)
- NOAA/Great Lakes Environmental Research Laboratory
- Great Lakes Observing System – Great Lakes Test-bed Working Group
- USDA/Agriculture Research Service, Pasture Systems and Watershed Management Research Unit
- Red River Basin Commission
- Upper Mississippi River Basin Association, Water Quality Program Director

- Illinois State Water Survey
- Minnesota Department of Agriculture
- Minnesota Discovery Farms
- University of Minnesota
- 2013 American Water Resources Association Specialty Conference on Agricultural Hydrology and Water Quality II

This list does not include the audiences of presentations conducted by other members of the Working Group over the last couple years. It was encouraging that many of these organizations immediately understood the RRAF as a best-available decision support tool that had limitations. Many expressed optimism and excitement about the possibility of a tool that could help reduce nutrient contamination in water bodies and improve overall water quality while addressing the largest uncontrollable factor in manure and fertilizer application: timing with respect to weather conditions. Many agencies were curious about the expansion timeline for the product.

A slow expansion of the RRAF has been in the planning stages at the NCRFC over the last year. It has been decided to follow two paths for introducing similar products to other areas. In the short-term, the same product and process developed in Wisconsin will be implemented in interested states. This means the same threshold techniques will be applied and no new validation will be done. Collaboration has begun with the Minnesota Department of Agriculture (DOA) and Minnesota Discovery Farms to set up a similar advisory webpage. Following this path, the invitation for other states within the NCRFC region to opt-in to the current setup does exist.

It should be noted that if another state agency within the NCRFC area wants to improve, refine, or modify their runoff risk guidance they are encouraged to do so. The NCRFC will happily provide historical model data for further validation and also provide real-time data for the forecast guidance. The NCRFC is hopeful for future collaboration opportunities with other agencies and universities to enhance the RRAF concept and product.

In long-term planning, the NCRFC has envisioned a second path for expanding the RRAF in the Midwest region. Planned for the summer of 2013, the NCRFC will begin to explore the applicability and capability of the National Weather Service's Office of Hydrologic Development (OHD) Hydrology Lab Research Distributed Hydrologic Model (HL-RDHM). This model will run the new Sacramento Heat Transfer and Evapotranspiration (SAC-HTET) model on a 4 km grid (Koren, et al., Oct 2010). It is hoped this newer model ran on a much finer scale will improve field scale simulations of surface runoff and improve the decision support tool being provided to manure producers. Figure 76 is an example image from the new model which highlights the boundaries and spatial resolution of the SAC-HTET model on the 4 km scale.

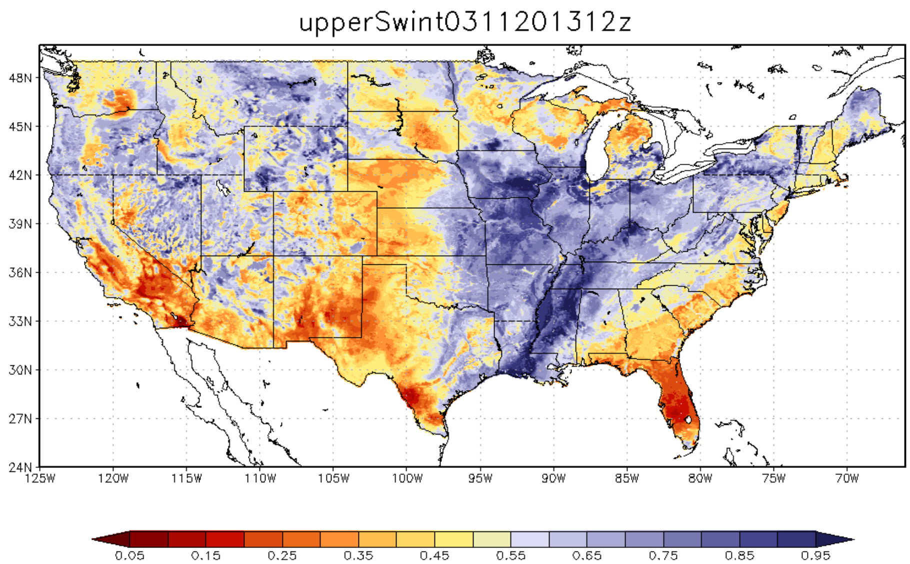


Figure 76. Example image of the SAC-HTET model ran on 4 km grids. This image represents the total soil moisture saturation averaged over the 0-25 cm layer.

5. CONCLUSIONS and RECOMMENDATIONS

The Runoff Risk Advisory Forecast (RRAF) is a first-of-its-kind decision support tool to aid Wisconsin farmers in making informed decisions about application timing of manure and fertilizers on their fields. Derived from land-atmosphere modeling at the North Central River Forecast Center (NCRFC), the RRAF incorporates the complex interaction of current and future soil conditions with future precipitation, snowmelt, and temperature forecasts to evaluate and categorize risk of runoff over the entire state. Farmers referencing the RRAF can gain some assurance in avoiding applying manure before runoff is predicted and therefore keep the valuable nutrients on their fields and out of nearby water bodies.

Overwhelming evidence in the last several decades has shown nutrients and other contaminates are often transported from agriculture fields into nearby water bodies. These contamination incidents can cause aquatic degradation on both the local scale (drinking well contamination, fish kills in local streams) as well as compounding large scale ecological issues such as hypoxic zones in the Gulf of Mexico and Great Lakes. Studies have shown that application timing before runoff occurs is a key factor in determining concentrations of contaminates carried off fields (Komiskey et al. 2011). The primary focus of the RRAF is to help manure producers manage the most challenging, and possibly most critical, aspect of manure management: deciding on when (or when not) to apply manure on their fields.

Decribing the livestock population of Wisconsin emphasizes the importance of a descision support tool such as the RRAF to help farmers safely manage the accompanying waste. At the end of 2012 the state was home to 3.5 million cattle (1.265 dairy cows), 3.5 million turkeys, 7 million chickens, and over 340,000 hogs and pigs (WDNR, 2012 [b]). While dairy production in Wisconsin ranks second in the country at generating over $4.5 billion in sales (USDA, 2011 [a]), it also leads in manure and waste production with an estimated 8 – 12 billion gallons annually (UW-Extension, 2012).

It is generally known Best Management Practices (BMPs) and Nutrient Management Plans (NMPs) do a decent job at helping farmers manage the "where" and "how" aspects of manure management. The goal of the RRAF is to augment this guidance with a more complete assessment of the future risk of runoff on their fields. Long-term success of the RRAF will depend on two factors: (1) will the RRAF be shown to be an accurate predictor of runoff risk for fields over time, and (2) will producers and spreaders build trust in the tool limiting application during risky periods and thus reducing the number of contamination incidents.

The development and implementation of the RRAF is a true success story of multi-agency collaboration at the local, state, and federal level. Many representatives provided guidance, suggestions, and data in the development process as well as remaining active participants of the RRAF Working Group. Model validation and verification was heavily based on edge-of-field (EOF) data provided by the Wisconsin Discovery Farms program and Wisconsin U. S. Geological Survey (USGS) office. Several departments of the University of Wisconsin were also involved. The primary agency at the state level is the Wisconsin Department of Agriculture, Trade, and Consumer Protection (DATCP) which takes the lead in owning, maintaining, and promoting the RRAF tool.

Besides addressing a major challenge in manure management, the RRAF fulfilled agency goals for both DATCP and the National Weather Service (NWS). Resulting from some highly damaging and publicized fish kill incidents in the mid-2000s, the state of Wisconsin convened a Manure Management Task Force (MMTF) which tasked DATCP to develop and implement a web-based runoff risk assessment and statewide notification system. The RRAF fits those requirements perfectly. As a decision support tool incorporating water quality ramifications, the RRAF also meets two future goals laid out by the National Weather Service in their Weather Ready Nation and vision (NWS, 2011).

The RRAF development and implementation followed the following process:

1. Defining simulated events from NCRFC models
2. Validating these events against observed EOF and small USGS watershed events
3. Generating basin specific thresholds to create two risk categories and reduce false alarm rates in the high risk category
4. Summarize and evaluate the spatial trends of daily RRAF products for 2011
5. Verify additional EOF observed runoff events against RRAF forecasts in 2011
6. Modify NCRFC models where needed to improve RRAF performance

The RRAF is made up of simulated runoff events derived from the Sacramento Soil Moisture Accounting Model (SAC-SMA) and Snow-17 model and is ran three times daily at the NCRFC. The model runs incorporate forecasts for precipitation and temperatures out five and ten days respectively. The time-series of three model states are evaluated to determine the presence of simulated runoff events. These are:

1. SAC-SMA Interflow runoff component > 0
2. SAC-SMA Upper Zone Tension Water Deficit = 0
3. Snow-17 Rain+melt >0

For each of the events simulated the interflow runoff is accumulated and compared against the basin threshold to categorize the event in either moderate (CAT 2) or high (CAT 3) risk. A list of events for the 216 basins in Wisconsin is sent to DATCP who populates the RRAF webpage with the derived basin risk for 72-hour periods.

The RRAF model was evaluated against observed data in two ways. A validation was performed by comparing observed runoff events against simulated events derived from historical model runs based on observed forcing data (precipitation and temperatures). This simulated dataset represents the model's best estimate of reality as uncertainty from forecast precipitation and temperatures is not included. The observed datasets used in the validation were derived from EOF scale and small USGS watersheds. Verification was the second evaluation where additional EOF events were compared against 2011 forecast runs of the RRAF where uncertainty in forcing data was a factor.

While evaluating the following summary of the validation and verification comparisons it should be stressed that perspective on the inherent process limitations needs to be applied. The most dramatic limitation with these comparisons is the scale difference between the lumped model basins and the observed datasets. The NWS basins in Wisconsin range between 9 and 1,837 mi^2 with an average size of 301 mi^2. The average EOF area was only 0.03 mi^2 (21.5 acres) while the selected USGS watersheds averaged about 15.9 mi^2 (10,200 acres). These scale differences result in an average of only 0.01% of the model basins covered by the EOF datasets and only 5% covered by the small USGS watersheds.

Alternatively, it could be stated that 95-99.99% of the NWS model basins are not considered when the RRAF was validated or verified.

Regardless of the scale difference, the validation indicated the RRAF model does a good job of detecting runoff events as 71% were captured in the moderate risk category (29% were missed) when averaging the EOF and USGS results. Applying the basin thresholds to focus only on high risk events showed 53% were hit by the model (47% were missed). However, the threshold dampened the false alarm rate from 58% for moderate risk to only 34% for high risk. Two concerns jump out from the results and need further discussion: (1) the high false alarms for moderate risk events (58%), and (2) the high percent of events missed for high risk events (47%).

Regarding the first concern, the ratio of median simulated hit and false alarm events for moderate risk was found to be over 8. This finding helped justify the use of thresholds to ignore the smallest events and remove many of these false alarms. This step created the high-risk category which focuses user attention on the larger magnitude events while reducing the false alarm rate as well. Applying the thresholds does reduce the percent of events hit (or Probability of detection, POD) in the high risk category which is alarming at first. However, going back to the median observed hit and miss event ratio also indicates a strong stratification exists with hits almost 8 times larger than the misses for the moderate risk category. With thresholds applied, the high risk ratio is over 9, which implies the events missed are essentially a magnitude smaller than the events that were hit. By removing the smaller events via the threshold, and lowering false alarms, no impact is seen on the capability of the model to continue capturing the most important larger runoff events.

Reviewing the Bias score for the combined validation is another beneficial method to evaluate the RRAF overall. Recalling that a perfect score for a model would be a Bias of 1.0, the moderate risk category recorded a score of 2.19. Applying the basin thresholds reduced false alarms substantially and provided a 94% improvement in the Bias score with a value of 0.93. The validation indicates that the RRAF model does a great job overall forecasting runoff risk for Wisconsin. This is especially apparent when the concern of false alarms is countered by the process limitation that only 95% or less of the model basin was accounted for by the observed datasets. The validation summary is described in section 3.14 and Table 18.

The next step in the RRAF development involved evaluating how the product behaved over time. A summary and analysis of the RRAF product in 2011 was accomplished by summarizing daily forecast runs for the year. Additional observed EOF data were used to complete verification on four of the NWS basins. As described in Chapter 4, for the simulated events in all of the 216 Wisconsin basins, 43% were moderate risk (CAT 2) and 57% were in the high-risk category (CAT 3). In addition, 95% of the total runoff simulated was in the high-risk events. Overall, 92% of the time no event was simulated in the basins, while moderate risk was represented 2% and high risk 6% of the time.

Another benefit of the 2011 analysis was the ability to evaluate the RRAF in Warning Days (WD) which represent 72-hour blocks of time and correspond to how risk is portrayed on the DATCP website. As described in sections 4.3.6 – 4.3.7, Warning Day 2 (forecast days 2 - 4) had the highest median percent of forecast runs with moderate, high, or any forecast risk with 15%, 23%, and 32% respectively.

When the simulated events in 2011 were broken down by runoff type, it was found 55% of the events were from rainfall only, 20% were from mixed events, and 25% were from snowmelt only. However, runoff itself was divided differently with mixed events accounting for 49% of the total runoff for a basin while rainfall only attributed 40% and snowmelt only just 11%. The analysis also indicated mixed rainfall-snowmelt events caused nearly half of the runoff in the high risk category.

The 2011 RRAF forecasts for four NWS basins were used in a verification analysis with observed runoff events from nine EOF sites. The same spatial scale limitation that was present in the validation exercise remained valid in this comparison. However, the uncertainty of using forecast rainfall and temperatures in the daily RRAF runs was added into this evaluation. It should also be noted that late fall and winter was a bit unusual for southern Wisconsin in 2011. It was warmer than normal during that time with dramatically less snowfall than normal.

Remarkably, the RRAF continued to do a great job capturing the occurrence of observed runoff events with a POD of 77% and 74% for moderate and high risk respectively. As with the validation, the misses that occurred were much smaller than the events categorized as hits. The false alarm rates were again a bit high with 66% and 54% for the moderate and high risk categories respectively. However, these values were very similar to rates found in the EOF validation (71% for moderate and 48% for high risk). No USGS data was available to average with the EOF values for the verification. Bias scores for moderate and high risk were calculated as 8.2 and 4.6 which are much poorer than the correlating validation scores. However, the verification consisted of very few observed hits and misses in relation to the number of simulated hits and false alarms due to the one year sample. Therefore the verification Bias scores are heavily skewed towards the simulated events and are not considered reliable. With that said, it still is a good sign that the high risk score does improve dramatically when compared to the moderate risk value.

As a result of the development process some changes were made to some NCRFC model parameters for a few basins that produced considerably different results when compared to their neighboring basins. Other future improvements consist of applying a second, lower threshold that will eliminate some of the smallest simulated events and reduce false alarms for the moderate risk category. Collaboration has begun with the Minnesota Department of Agriculture (DOA) to develop a similar RRAF product for the state of Minnesota. The NCRFC also plans to start evaluating a new 4-km distributed SAC-HTET model for determining if a better tool can be produced which will allow easier expansion across the upper Midwest.

Overall, the results of the RRAF validation and verification appear to be more than satisfactory for a first generation decision support tool for manure producers and farmers in Wisconsin. The results appear even more impressive when the spatial scale limitations are taken into consideration. Encouraging feedback from manure producers and farmer advocates as well as from state and local conservation and regulatory agencies emphasizes both the need and demand for this type of product and guidance. It is not unreasonable to propose that consistent monitoring of the RRAF by farmers in their day-to-day manure, and fertilizer, management decisions could produce advantages across the spectrum. More manure and nutrients would remain on the fields maximizing the nutrient value and crop yields for farmers. Likewise, fewer nutrients transported from the fields would decrease aquatic degradation at both the local level (streams and rivers) as well as at the regional and national scales (Great Lakes, Gulf of Mexico). This dual-benefit outcome easily supports why the National Weather Service, DATCP, and the rest of the Working

Group will continue to support and improve the RRAF decision support tool into the future.

6. REFERENCES

Anderson, E. A. (1976). *A point energy and mass balance model of a snow cover.* NOAA Technical Report NWS 19.

Anderson, E. A. (2006). *Snow accumulation and ablation model - Snow-17, NWSRFS user manual.* NWS Hydrology Laboratory. Retrieved from http://www.nws.noaa.gov/oh/hrl/nwsrfs/users_manual/part2/_pdf/22snow17.pdf

Brenneman, G. (n.d.). *Iowa Manure Manager Series Volume 3: Winter Application*. Retrieved December 2012, from Iowa State University Extension: http://www.agronext.iastate.edu/immag/pubs/imms/vol3.pdf

Burnash, R. C. (1995). The NWS river forecast system - Catchment modeling. In V. Singh, *Computer Models of Watershed Hydrology* (pp. 311-366).

DATCP. (2011, November). *Wisconsin Nutrient Management Update.* Retrieved December 2012, from Wisconsin Department of Agriculture, Trade, and Consumer Protection: http://datcp.wi.gov/uploads/Farms/pdf/Update2011News.pdf

DATCP. (2012 [a]). *Frequently Asked Questions Nutrient Management Rule.* Retrieved December 2012, from Wisconsin Department of Agriculture, Trade, and Consumer Protection: http://datcp.wi.gov/uploads/Farms/pdf/FrequentlyAskedQA.pdf

DATCP. (2012 [b], November). *Wisconsin Nutrient Management Update.* Retrieved December 2012, from Wisconsin Department of Agriculture, Trade, and Consumer Protection: http://datcp.wi.gov/uploads/Farms/pdf/2012NutrientMgmtNews.pdf

Dickrell, J. (2009, October 2). *Make Manure Pay*. Retrieved December 2012, from AGWEB: http://www.agweb.com/article/Make_Manure_Pay_196551/

Discovery Farms. (2011, September). *Management Considerations for Late Summer Manure Application.* Retrieved December 2012, from University of Wisconsin Discovery Farms: http://www.uwdiscoveryfarms.org/Publications/ViewAllPublications.aspx

Discovery Farms. (2012 [a]). *About UW Discovery Farms*. Retrieved December 2012, from University of Wisconsin Discovery Farms: http://www.uwdiscoveryfarms.org/AboutUs.aspx

Discovery Farms. (2012 [b]). *Nutrient Management*. Retrieved December 2012, from University of Wisconsin Discovery Farms: http://www.uwdiscoveryfarms.org/Education/NutrientManagement.aspx

Discovery Farms. (n.d.). *Runoff Lessons: Frozen and Snow-Covered Ground.* Retrieved December 2012, from University of Wisconsin Discovery Farms: http://www.uwdiscoveryfarms.org/Publications/ViewAllPublications.aspx

Farm Journal. (2010, May 15). *Ask an Agronomist*. Retrieved December 2012, from AGWEB: http://www.agweb.com/blog/ask_an_agronomist_252/Manure_Versus_Commercial_Fertilizers_15704/

Fleming, R., & Fraser, H. (2000). *Impacts of Winter Spreading of Manure on Water Quality - Literature Review.* Retrieved December 2012, from http://www.ridgetownc.uoguelph.ca/research/documents/fleming_manure.pdf

Frame, D. (2010, August 2). *When does runoff occur and what can be done to reduce impacts on water quality?* Retrieved December 2012, from University of Wisconsin Discovery Farms: http://www.slideshare.net/aradatz/non-frozen

Frame, D. (2011, August 31). *Manure Spreading: Things to Consider Part2!* Retrieved December 2012, from University of Wisconsin Extension: http://fyi.uwex.edu/discoveryfarms/2011/08/manure-spreading-things-to-consider-part-2/

Gauldin, C. (2008, August 25). *Study finds manure's economic value as fertilizer replacement*. Retrieved December 2012, from Feedstuffs: http://feedstuffs.com/ME2/dirmod.asp?sid=&nm=&type=Publishing&mod=Publications%3A%3AArticle&mid=AA01E1C62E954234AA0052ECD5818EF4&tier=4&id=D7E9D97DB3FE4EAA85DEA6EF2BB1D602

Good, L. (2012, November 29). email communication.

Jenks, M. (2012, November 29). email communication.

Komiskey, M., Stuntebeck, T., Frame, D., & Madison, F. (2011). Nutrients and sediment in frozen-ground runoff from no-till fields receiving liquid-dairy and solid-beef manures. *Journal of Soil and Water Conservation, 66*(5).

Koren, V., Smith, M., Cui, Z., Cosgrove, B., Werner, K., & Zamora, R. (Oct 2010). *Modification of Sacramento Soil Moisture Accounting Heat Transfer Component (SAC-HT) for Enhanced Evapotranspiration.* NOAA, National Weather Service. U.S. Department of Commerce.

Lorimor, J. (1995). *Fate of Nutrients from Liquid Swine Manure Land Applied in the Winter.* Ph.D. dissertation Iowa State University Ames, IA.

McClure, S. (2009, August 25). Baseflow and event separation methodology using USGS daily streamflow data. *Memorandum from DATCP*.

Mississippi River Gulf of Mexico Watershed Nutrient Task Force. (2008). *Gulf Hypoxia Action Plan 2008.* Retrieved from http://water.epa.gov/type/watersheds/named/msbasin/upload/2008_8_28_msbasin_ghap2008_update082608.pdf

Mississippi River Gulf of Mexico Watershed Nutrient Task Force. (2011). *FY 2011 Annual Report.* Retrieved from http://water.epa.gov/type/watersheds/named/msbasin/upload/Hypoxia_Task_Force_Annual_Report_2011.pdf

NRCS. (2005). *Wisconsin Code 590 Nutrient Management Conservation Practice Standard.* U.S. Natural Resources Conservation Service. Retrieved from http://efotg.sc.egov.usda.gov/references/public/WI/590.pdf

NWS. (2011). *Weather-Ready Nation: NWS Strategic Plan 2011.* National Weather Service. Retrieved from http://www.nws.noaa.gov/com/weatherreadynation/files/strategic_plan.pdf

NWS. (2012, December 5). *Dual-Polarization Radar Training for NWS Partners*. Retrieved from National Weather Service Warning Decision Training Branch: http://www.wdtb.noaa.gov/courses/dualpol/Outreach/index.html

Stuntebeck, T. D., Komiskey, M. J., Owens, D. W., & Hall, D. W. (2008). Methods of data collection, sample processing, and data analysis for edge-of-field, streamgaging, subsurface-tile, and meteorological stations at Discovery

Farms and Pioneer Farm in Wisconsin, 2001-7. *U.S. Geological Survey Open-File Report 2008-1015*.

Stuntebeck, T. D., Komiskey, M. J., Peppler, M. C., Owens, D. W., & Frame, D. R. (2011). Precipitation-runoff relations and water-quality characteristics at edge-of-field stations, Discovery Farms and Pioneer Farm, Wisconsin, 2003-8. *U.S. Geological Survey Scientific Investigations Report 2011-5008*.

UM-Extension. (2010, October 19). *UW-Extension Agents and Specialists Earn Award for their Innovate Education on Handling Manure*. Retrieved December 2012, from University of Minnesota Extension: http://www.extension.org/pages/30659/uw-extension-agents-and-specialists-earn-award-for-their-innovate-education-on-handling-manure

UM-Extension. (2011 [a], February 14). *Environmental Benefits of Manure Application*. Retrieved December 2012, from University of Minnesota Extension: http://www.extension.org/pages/14879/environmental-benefits-of-manure-application

UM-Extension. (2011 [b], February 14). *Preferential Flow of Manure in Tile Drainage*. Retrieved December 2012, from University of Minnesota Extension: http://www.extension.org/pages/27488/preferential-flow-of-manure-in-tile-drainage

UM-Extension. (2012 [a], March 7). *Environmental Impacts and Benefits of Manure: Phosphorus and Surface Water Protection*. Retrieved December 2012, from University of Minnesota Extension: http://www.extension.org/pages/19859/environmental-impacts-and-benefits-of-manure:-phosphorous-and-surface-water-protection

UM-Extension. (2012 [b], June 26). Introduction to Agriculture - The Minnesota Dairy Industry. *University of Minnesota Extension*. Retrieved December 2012, from http://www.youtube.com/watch?feature=player_embedded&v=2n7BBK3uzTA

UM-Extension. (2012 [c], April 28). *Manure Application Considerations for this Spring*. Retrieved December 2012, from University of Minnesota Extension: http://www1.extension.umn.edu/dairy/manure/manure-application-considerations/

UM-Extension. (2012 [d]). *What's Manure Worth?* Retrieved December 2012, from University of Minnesota Extension: http://www1.extension.umn.edu/agriculture/manure-management-and-air-quality/manure-application/whats-manure-worth/

USDA. (1992). *Agriculture Waste Management Field Handbook.* U.S. Department of Agriculture. Retrieved December 2012

USDA. (2011 [a]). *2011 Wisconsin State Agriculture Overview.* Retrieved December 2012, from U.S. Department of Agriculture - National Agriculture Statistics Service: http://www.nass.usda.gov/Statistics_by_State/Ag_Overview/AgOverview_WI.pdf

USDA. (2011 [b]). *Milk Cows: Production per Cow and Total Milk Production, by County, Wisconsin, 2010-2011.* Retrieved December 2012, from U.S. Department of Agriculture - National Agriculture Statistics Service: http://www.nass.usda.gov/Statistics_by_State/Wisconsin/Publications/County_Estimates/milk_cows.pdf

USEPA. (1990). *National Survey of Pesticides in Drinking Water Wells, Phase I Report. EPA-570990015.* U.S. Envirionmental Protection Agency.

USEPA. (2000). *National Water Quality Inventory.* U.S. Environmental Protection Agency. Retrieved from http://water.epa.gov/lawsregs/guidance/cwa/305b/2000report_index.cfm

USEPA. (2002). *Environmental and Economic Benefit Analysis of Final Revisions to the National Pollutant Discharge Elimination System Regulation and the Effluent Guidelines for Concentrated Animal Feeding Operations.* U.S. Environmental Protection Agency. Retrieved December 2012, from http://www.epa.gov/npdes/pubs/cafo_benefit_p1.pdf

USEPA. (2012 [a]). *Hypoxia 101*. Retrieved December 2012, from Mississippi River Gulf of Mexico Watershed Nutrient Task Force: http://water.epa.gov/type/watersheds/named/msbasin/hypoxia101.cfm

USEPA. (2012 [b]). *NPDES Permit Writers' Manual for Concentrated Animal Feeding Operations.* U.S. Environmental Protection Agency. Retrieved December 2012, from http://www.epa.gov/npdes/pubs/cafo_permitmanual_entire.pdf

UW-Extension. (2009). *2009 MMS Speakers and Bios*. Retrieved December 2012, from University of Wisconsin Extension Midwest Manure Summit: http://fyi.uwex.edu/midwestmanure/previous-summits/2009-midwest-manure-summit/2009-mms-speakers-and-bios/

UW-Extension. (2011 [a], March 16). *How Much Water am I Adding to that Field?* Retrieved December 2012, from University of Wisconsin Extension: http://fyi.uwex.edu/discoveryfarms/2011/03/how-much-water-am-i-adding-to-that-field/

UW-Extension. (2011 [b], June 28). *The Importance of 'That One Big Storm'*. Retrieved December 2012, from University of Wisconsin Extension: http://fyi.uwex.edu/discoveryfarms/2011/06/the-importance-of-that-one-big-storm/

UW-Extension. (2012). *Nutrient Managment Fast Facts.* Retrieved December 2012, from University of Wisconsin Extension: http://ipcm.wisc.edu/download/pubsNM/NutrientManagementFastFacts.pdf

Wahl, T. L., & Wahl, K. L. (2007, April 13). *BFI: A computer program for determining an index to base flow*. Retrieved from U.S. Bureau of Reclamation: http://www.usbr.gov/pmts/hydraulics_lab/twahl/bfi/

WDNR & DATCP. (2006). *Final Report Manure Management Task Force.* Wisconsin Department of Natural Resoureces & Department of Agriculture, Trade, and Consumer Protection. Retrieved December 2012, from http://www.wistatedocuments.org/cdm/ref/collection/p267601coll4/id/183

WDNR. (2004). *Wisconsin's Runoff Rules: What Farmers Need to know.* Wisconsin Department of Natural Resources. Retrieved from http://dnr.wi.gov/topic/nonpoint/documents/farmersneed.pdf

WDNR. (2011, November). *Wisconsin's WPDES Permitted Concentrated Animal Feeding Operations.* Retrieved December 2012, from Wisconsin Department of Natural Resources: http://dnr.wi.gov/topic/AgBusiness/documents/cafo_map.pdf

WDNR. (2012 [a], December 4). *Agri-Business and CAFOs*. Retrieved December 2012, from Wisconsin Department of Natural Resources: http://dnr.wi.gov/topic/AgBusiness/

WDNR. (2012 [b], November 14). *Concentrated Animal Feeding Operations*. Retrieved December 2012, from Wisconsin Department of Natural Resources: http://dnr.wi.gov/topic/AgBusiness/CAFOgeneralInfo.html

Wisconsin Administrative Code. (2007). *Chapter Natural Resources (NR-243) Animal Feeding Operations.* Retrieved from http://docs.legis.wisconsin.gov/code/admin_code/nr/243.pdf

Wisconsin Administrative Code. (2011). *Chapter Agriculture, Trade, and Consumer Protection (ATCP-50) Soil and Water Resource Management Program.* Retrieved from http://docs.legis.wisconsin.gov/code/admin_code/atcp/50.pdf

Wisconsin Administrative Code. (2012). *Chapter Natural Resources (NR-151) Runoff Management.* Retrieved from http://docs.legis.wi.gov/code/admin_code/nr/151.pdf