Reviewer 1 Comments and Response:

Compared to the previous version, this updated version of the manuscript has improved a lot. Although I should note that not all suggestions raised have been accounted for. However, the performed work, results and conclusions are well presented.
There is only one final detail that I would like to see altered before I feel the manuscript is ready for publication. At multiple places (both in the abstract and conclusion) the authors talk about 1100 of radar precipitation observations. However, this is just 46 days of data. At another location details are provided that actually 400 of the 1100 hours contained precipitation. I would therefore suggest that the authors alter the 1100 into 400 hours of precipitation

We thank the reviewer for the above comments. We have updated the necessary changes throughout the manuscript to properly reflect the correct number of days and precipitation amounts.

Formatted: Font color: Text 1

Formatted: Font: 10 pt, Bold, Font color: Text 1

Reviewer 2 Comments and Responses:

The paper could provide a long-term verification of dual-pol QPE algorithms which is relevant for hydrology. The authors stress that they focus on the range effect but this is in contradiction with the extended list of objectives in the introduction and the limited amount of results related to range in the conclusions

Formatted: Font: 12 pt, Bold

We appreciate this comment. We have added discussion in the conclusions to elaborate upon this aspect. We also added elaboration in the list of objects (near line 79 on page 3) to emphasize the range effect.

The number of data is limited. Why only one year? Why only 46 days of precipitation are available when the normal is around 100 days?

Formatted: Font: 12 pt, Bold

We chose a random year for the analyses to be conducted, we elaborated that 100 days have 'measureable' rainfall (i.e., greater than trace) whereas 50 days have greater than 0.5mm in of rainfall. Therefore, the 46 days chosen / analyzed falls near the average amount of days with appreciable rainfall.

The conclusions are short and do not summarize clearly the main findings (i.e. the algorithm's relative performance in function of the range). A proper discussion on the validity and possible cause of the different results is missing

Formatted: Font: 12 pt, Bold

Thank you for this comment. We have expanded upon the conclusions which were lacking in wrapping the paper up.

The information on the data and their quality is still limited while it seems some observation errors affect the results. Which type of quality control is effectively performed by WDSS-II on the radar data? Why not using the one-hour precipitation product of NOAA as reference? Why using the Mesonet network when the higher resolution CoCoRaHS is considered as better by the authors? The data selection criteria and choice of statistics are not sufficiently discussed.

Formatted: Font: 12 pt, Bold

We have added a more detailed description of the quality controlled techniques implemented, which would mitigate large errors in QPE from various modules within the WDSS-II framework. We did not consider using the DP rate as a reference, as that is more of a heuristic algorithm that blends multiple different algorithms together (it is difficult to determine whether they implement R(KDP), R(Z,ZDR), etc.) without doing a deep analysis of the radar parameter values as well as the particular algorithm implemented at each time. Furthermore, it is difficult to determine whether each of the 3 radar locations implement the same sort of dual-pol radar equation at the same times. Lastly, We implemented the Mesonet data due to the timing in which the current study was conducted. The authors have follow-up studies which utilize CoCoRaHS, HADS, MADIS, and other gauge locations.

In Figure 2, the results vary a lot between the algorithm's and the radars making interpretations difficult. I am surprised by the bad performance of KDP (did you check the cause visually?). The tentative explanations of radar issues for specific gauges (e.g. bright band effect) are not robust. In Figure 3-8, only the overall best and worst algorithm's are shown, which is too limited (I would present the best of each type). It is often unclear for which algorithm an interpretation is valid.

Formatted: Font: 12 pt, Bold

We thank the reviewer for these comments. After checking visually, bright-banding were present in several cases, but the w2qcnndp as well as w2qcnn algorithms *should* have handled them effectively (cases slip through, of course). We have addressed this within the text which is primarily the result of the large biases observed in spite of larger distance from the radar. The algorithms represented via Figures 3-8 are labeled within the caption and represent the best-performing R(Z,ZDR) and worst performing R(ZDR,KDP) algorithms. This helps to highlight differences between the algorithms not only between the warm, but also the cool season.

The results of similar studies (including from the authors) are not properly reviewed. Is there a connection with your recently submitted article on X-Band?

Formatted: Font: 12 pt, Bold

We have seen similarities with the superiority of R(Z,ZDR) algorithms over R(ZDR,KDP) or R(KDP). We did, as well, see superiority with the R(Z)-Convective equation as well.

The description of the statistical analyses needs to be much more clear and precise (proper definition and interpretation, thresholds used for zeros, selection of hit only data)

Formatted: Font: 12 pt, Bold

We appreciate this comment, and have elaborated on the definition of thresholds and hit only at the end of the statistical analyses section in which more than 2 tips were needed for calculations to be implemented.

The new title sounds a bit odd to me

Formatted: Font: 12 pt, Bold

We have changed the title of the article to make it flow better.

The abstract has not been improved as suggested and is not consistent with the conclusions.

Formatted: Font: 12 pt, Bold

We appreciate this reviewer comment, and have expanded on the abstract to better reflect the conclusion, make it easier to read, and fixed some spelling errors.

The comments have not been taken into account. There is still part of the methodology in the "results" section.

Formatted: Font: 12 pt, Bold

We thank the reviewer for this comment, and we have moved text to/from the methodology and results section to better reflect the text within each section.

No significant efforts have been made to improve the text structure, terminology and style. There are annoying editing errors at this stage (e.g. a repeated sentence on line 212)

Formatted: Font: 12 pt, Bold

We have moved text around throughout the methodology and results to create a better-flowing manuscript.

Some definitions are still incorrect or imprecise

The authors thank the reviewer for this comment. We have gone through the text and ensured accuracy in the definition and spelling of each acronym.

The results section is still not clear nor concise. There are too much points in Figure 2. There are too much plots in the figures. I would show only NMB, NME, PoFD, PoD. Paragraphs over the different radars could be combined. What is exactly on figures 2-8 : best at each point (your response) or only R(Z,ZDR) (figure caption)?

We have utilized only the best algorithm from the set of R(Z,ZDR) equations and the worst algorithm from the set of R(ZDR,KDP) equations as these consistently showed to be the best and worst, respectively. We have implemented the data from each of the statistical analyses to better represent the performance of each algorithm at each radar. Some results would not have been accounted for or even could have been completely missed without some of the statistical measures analyzed in this fashion.

The number and quality of the references are acceptable but they are often cited for anecdotal reasons (e.g. Figueras et al. on line 381). They are best used for discussion in the introduction and conclusions sections.

We have altered our references and moved them around to be more appropriate for the current study.

Formatted: Font: 12 pt, Bold

Formatted: Font: 12 pt, Bold

Formatted: Font: 12 pt, Bold

~~RANGE AS A FUNCTION OF~~ DUAL-POLARIZED QUANTITATIVE PRECIPITATION ESTIMATION AS A FUNCTION OF RANGE

Micheal J. Simpson[1] and Neil I. Fox[2]

[1]University of Missouri, School of Natural Resources, Water Resources Program, Department of Soil, Environmental, and Atmospheric Sciences, 203-T ABNR Building, Columbia, Missouri, USA, 65211. Tel: +001 4053256459 Email: mjs5h7@mail.missouri.edu

~~[1]Cooperative Institute of Mesoscale Meteorological Studies, University of Oklahoma. National Severe Storms Laboratory, Norman, Oklahoma. Tel: +001 4053256459. Email: micheal.simpson@noaa.gov~~

[2]University of Missouri, School of Natural Resources, Water Resources Program, Department of Soil, Environmental, and Atmospheric Sciences, 332 ABNR Building, Columbia, Missouri, USA, 65211. Tel: +001 5738822144 Email: FoxN@Missouri.edu

Correspondence to: Micheal J. Simpson (micheal.simpson@noaa.gov)

Abstract. Since the advent of dual-polarization radar technology, many studies have been conducted to determine the extent to which the differential reflectivity (ZDR) and specific differential phase shift (KDP) add benefits to estimating rain rates (R) compared to reflectivity (Z) alone. It has been previously noted that this new technology provides significant improvement to rain rate estimation, primarily for ranges within 125 km of the radar. Beyond this range, it is unclear as to whether the National Weather Service conventional R(Z)-Convective algorithm is superior, as little research has investigated radar precipitation estimate performance at larger ranges. The current study investigates the performance of three radars, St. Louis (KLSX), Kansas City (KEAX), and Springfield (KSGF), MO, with 15 tipping bucket gauges serving as ground-truth to the radars. With over 300 hours of precipitation data were analyzed for the current study, it was found that, in general, performance degraded with range beyond,

Formatted: Font: (Default) Times New Roman

Formatted: Normal, No bullets or numbering

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman approximately, 150 km from each of the radars. Probability of detection in addition to bias values decreased, while the false alarm rates increased as range increased. Bright-band contamination was observed to play a potential role as large increases in the absolute bias and overall error values near 120 km for the cool season, and 150 km in the warm season. Furthermore, upwards of 60% of the total error was due to precipitation falsely estimated, while 20% of the total error was due to missed precipitation. Correlation coefficient values increased by as much as 0.4 when these instances were removed from the analyses (i.e., hits only). Overall, due to the lowest normalized standard error of less than 1.0, a National Severe Storms Laboratory (NSSL) R(Z,ZDR) equation was determined to be the most robust, while a R(ZDR,KDP) algorithm recorded NSE values as much as 5. The addition of dual-polarized technology was shown to better estimate quantitative precipitation estimates than the conventional equation. The analyses further our understanding in the strengths and limitations of the Next Generation Radar system overall, and from a seasonal perspective.

Formatted: Indent: Left: 0.5", No bullets or

~~**Abstract.** Since the advent of dual-polarized~~ation radar ~~technology, many studies have been conducted to determine the extent to which the differential reflectivity (ZDR) and specific differential phase shift (KDP) add benefits to estimating rain rates (R)~~ compared ~~to reflectivity (Z). It has been previously noted that this new technology provides significant improvement to rain rate estimation, but only for ranges within 125 km from~~ of ~~the radar. Beyond this range, it is unclear as to whether the National Weather Service conventional R(Z)-Convective algorithm is superior, as little research has investigated radar precipitation estimate performance at large ranges. The current study investigates the performance of three radars, St. Louis (KLSX), Kansas City (KEAX), and Springfield (KSGF), MO, with respect to range, with 15 terrestrial-based tipping bucket gauges served~~ serving ~~as ground-truth to the radars. Over 1100~~ 300 ~~hours of precipitation data were analyzed for the current study. It was found that, in general, performance degraded with range beyond, approximately, 150 km from the radar. Probability of detection in addition to bias values decreased, while the false alarm rat~~e~~ios increased as range increased. Bright-~~

Formatted: Highlight

Formatted: Highlight

~~band contamination was observed to play a potential role as large increases in the absolute bias and overall error values near 120 km for the cool season, and 150 km in the warm season. Addition of dual-polarized technology was shown to better estimate quantitative precipitation estimates than the conventional equation. The analyses found further our understanding in the strengths and limitations of the Next Generation Radar system overall, and from a seasonal perspective.~~

1 Introduction

In 2012, the National Weather Service (NWS) began upgrading the Next Generation Radar (NEXRAD) system from single- to dual-polarization. The potential benefits of this upgrade were investigated by the National Severe Storms Laboratory (NSSL) and the Cooperative Institute for Mesoscale Meteorological Studies. These advantages include, but are not limited to, (1) significant improvement in radar rainfall estimation (Ryzhkov et al., 2005; Gourley et al., 2010) through better representation of precipitation shape (Brandes et al., 2002; Gorgucci et al., 2000, 2006), (2) discrimination between solid and liquid precipitation (Zrnic and Ryzhkov, 1996), allowing for better distinction between areas of heavy rain and hail (Park et al., 2009; Giangrande and Ryzhkov, 2008; Cunha et al., 2013), (3) identifying the melting layer position in the radar field (Straka et al., 2000; Park et al., 2009), and (4) calculating drop-size distributions retrieved from measurements of reflectivity (Z), differential reflectivity (ZDR), and specific differential phase shift (KDP) as opposed to using ground-based point located disdrometers (Zhang et al., 2001; Brandes et al., 2004; Anagnostou et al., 2008).

Rain rate retrieval by weather radars is an estimation based upon the dielectric properties of the hydrometeors encountered in the atmosphere. Therefore, there is no direct measurement of rainfall, and this inherently introduces error. However, dual-polarized radar technology allows for in-depth analyses on the microphysics of precipitation that single-polarization was incapable of conducting. In spite of this technology, conflicting studies report the benefits for quantitative precipitation estimation (QPE). For example, Gourley et al. (2010) and Cunha et al. (2015) reported that conventional R(Z) algorithms have significantly better bias than algorithms containing ZDR and/or KDP, while others (e.g., Ryzhkov et al., 2013; Simpson et al., 2016) report the opposite. This could be due, at least in part, to the fact that hydrometeor types (e.g., rain versus hail) vary on spatial scales that cannot be easily resolved by even densely gauged networks.

Multiple studies have found that the performance of radar rain rate estimates decrease as range increases (Smith et al., 1996; Ryzhkov et al., 2003) which is caused, primarily, by degradation of beam quality with range. Furthermore, the researchers also discuss how the probability of detection at larger ranges decreases, as the radar beam overshoots shallow, stratiform precipitation, especially winter precipitation. Bright-banding can also play a crucial role in significantly increasing the amount of precipitation estimated by the radar, prompting many researchers to produce automated bright-band detection algorithms (e.g., Zhang et al., 2008; . Zhang and Qi, 2010).

Despite these overall disadvantages, studies have shown that radar rainrate algorithms seldom exceed absolute errors on the order of 10 mm h^{-1}. However, many of these studies have looked at a small sample of rain events (on the order of 10-50 hours) (Kitchen and Jackson, 1993; Smith et al., 1996; Ryzhkov et al., 2003; Gourley et al., 2010; Cunha et al., 2013). Long-term performances of weather radar are becoming more common in recent years as the availability of data becomes more abundant (e.g., Haylock et al., 2008; Goudenhoofdt and Delobbe, 2012; Fairman et al., 2015; Goudenhoofdt and Delobbe, 2015). Additionally, few studies (e.g., Smith et al., 1996; Cunha et al., 2015; Simpson et al., 2016) quantified QPE errors including the probability of detection and false alarm ratio. In order to gain a better understanding of the performance of weather radars on rain rate estimates, more data must be collected over a broad range of precipitation regimes in addition to an overall broader region of interest.

The overarching objective of the current study was to assess the performance of three different radars within the state of Missouri at various ranges from the radar, using terrestrial-based tipping bucket gauges as ground-truth data. Radar rain rate estimation algorithms include 55 algorithms encompassing standard R(Z) relations as well as algorithms containing dual-polarization variables including differential reflectivity (ZDR) and the specific differential phase shift (KDP). A rain rate echo classification algorithm was also tested for performance in correctly identifying the suitable rain rate algorithm to choose based on the Z, ZDR, and KDP radar fields. The current work expands upon that of Simpson et al. (2016) such that a larger sample of data was analyzed (over 300 hours of rainfall data from forty-six separate days in 2014) to encompass multiple different precipitation regimes for both summer and winter, with several ground-truth tipping buckets to analyze the performance of three separate radars as a function of range, and further expanding upon the effects of erroneous precipitation estimates on the overall radar error. Objectives for this study included, (1) statistically analyze the performance of each radar at various ranges (compared against the gauges), (2) compute (a) the amount of precipitation incorrectly estimated by the radar (quantifying the probability of false detection) and (b) the amount of precipitation incorrectly missed by the radar but measured by the rain gauge, (3) test the overall best radar rain rate algorithm, and (4) perform objectives (1), (2), and (3) while the data is separated into warm and cool seasons which have been shown to result in significantly different QPE's (Smith et al., 1996; Ryzhkov et al., 2003; Cunha et al., 2015).

~~In 2012, the National Weather Service (NWS) began upgrading the Next Generation Radar (NEXRAD) system from single- to dual-polarization. The potential benefits of this upgrade were investigated by the National Severe Storms Laboratory (NSSL) and the Cooperative Institute for Mesoscale Meteorological Studies. These advantages include, but are not limited to, (1) significant improvement in radar rainfall estimation (Ryzhkov et al., 2005; Gourley et al., 2010) through better representation of precipitation shape (Brandes et al., 2002; Gorgucci et al., 2000, 2006), (2) discrimination between solid and liquid precipitation (Zrnic and Ryzhkov, 1996), allowing for better distinction between areas of heavy rain and hail (Park et al., 2009; Giangrande and Ryzhkov, 2008; Cunha et al., 2013), (3) identifying the melting layer position in the radar field (Straka et al., 2000; Park et al., 2009), and (4) calculating drop-size distributions retrieved from measurements of reflectivity (Z), differential reflectivity (ZDR), and specific differential phase shift (KDP) as opposed to using ground-based point located disdrometers (Zhang et al., 2001; Brandes et al., 2004; Anagnostou et al., 2008).~~

~~Rain rate retrieval by weather radars is an estimation based upon the dielectric properties of the hydrometeors encountered in the atmosphere. Therefore, there is no direct measurement of rainfall, and this inherently introduces error. However, dual-polarized radar technology allows for in-depth analyses on the microphysics of precipitation that single-polarization was incapable of conducting. In spite of this technology, conflicting studies report the benefits for quantitative precipitation estimation (QPE). For example, Gourley et al. (2010) and Cunha et al. (2015) reported that conventional R(Z) algorithms have significantly better bias than algorithms containing ZDR and/or KDP, while others (e.g., Ryzhkov et al., 2013; Simpson et al., 2016) report the opposite. This could be due, at least in part, to the fact that hydrometeor types (e.g., rain versus hail) vary on spatial scales that cannot be easily resolved by even densely gauged networks.~~

~~Multiple studies have found that, in general, the performance of radar rain rate estimates decrease as range increases (Smith et al., 1996; Ryzhkov et al., 2003) which is caused, primarily, by degradation of beam quality and broadening of the beam with range. Furthermore, the researchers also discuss how the probability of detection at larger ranges decreases, as the radar beam overshoots shallow, stratiform precipitation, including winter storms. Bright-banding can also play a crucial role in significantly increasing the amount of precipitation estimated by the radar.~~

~~Despite these overall disadvantages, studies have shown that radar rainrate algorithms seldom exceed absolute errors on the order of 10 mm h~~$^{-1}$~~. However, many of these studies have looked at a small sample of rain events (on the order of 10-50 hours) (Kitchen and Jackson, 1993; Smith et al., 1996; Ryzhkov et al., 2003; Gourley et al., 2010; Cunha et al., 2013). Long-term performances of weather radar are becoming more common in recent years as the availability of data becomes more abundant (e.g., Haylock et al., 2008; Goudenhoofdt and Delobbe, 2012; Fairman et al., 2015; Goudenhoofdt and Delobbe, 2015). Additionally, few studies (e.g., Smith et al., 1996; Cunha et al., 2015; Simpson et al., 2016) quantified meteorologically significant measures including the probability of detection and false alarm ratio. In order to get a better understanding of the performance of weather radars on rain rate~~

~~estimates, more data must be collected over a broad range of precipitation regimes in addition to an overall broader region of interest.~~

~~The overarching objective of the current study was to assess the overall performance of three different radars within the state of Missouri at various ranges from the radar, using terrestrial-based tipping bucket gauges as ground-truth data. Radar rain rate estimation algorithms include 55 algorithms encompassing standard R(Z) relations, in addition to~~and ~~algorithms containing dual-polarization variables including ZDR and KDP. A rain rate echo classification algorithm was also tested for performance in correctly identifying the suitable rain rate algorithm to choose based on the Z, ZDR, and KDP radar fields. The current work expands upon that of Simpson et al. (2016) such that a larger sample of data were~~ was ~~analyzed (over 10300 hours of rainfall data from forty-six separate days in 2014) to encompass multiple different precipitation regimes for both summer and winter, with several ground-truth tipping buckets to analyze the performance of three separate radars as a function of range, and further expanding upon the effects of erroneous precipitation estimates on the overall radar error. Objectives for this study included, (1) statistically analyze the performance of each radar at various ranges (compared against the terrestrial-based gauges), (2) compute (a) the amount of precipitation incorrectly estimated by the radar (quantifying the probability of false detection) and (b) the amount of precipitation incorrectly missed by the radar but measured by the rain gauge, (3) test the overall best radar rain rate algorithm, and (4) perform objectives (1), (2), and (3) while the data is separated into warm and cool seasons which have been shown to result in significantly different QPE's (Smith et al., 1996; Ryzhkov et al., 2003; Cunha et al., 2015).~~

Formatted: Highlight

Formatted: Highlight

2 Study area and methods

2.1 Study area

National Weather Service radars from St. Louis (KLSX), Kansas City (KEAX), and Springfield (KSGF), MO are able to scan the majority of the state of Missouri. Because of this, the three aforementioned radars were used to assess overall performance in estimating precipitation for this study. Each radar covered a 200-km radius for which a different number of gauges were within the domain: KLSX, KEAX, and KSGF covered 9, 8, and 5 gauges, respectively (Figure 1).

Missouri is characterized as a continental type of climate, marked by relatively strong seasonality. Furthermore, Missouri is subject to frequent changes in temperature, primarily due to its inland location and its lack of proximity to any large lakes. All of Missouri experiences below-freezing temperatures on a yearly-basis. For example, the majority of the state typically registers, 110 days with temperatures below freezing, while the Bootheel (i.e., southeast region) records, on average, 70 days of below freezing day temperatures, emphasizing the typical northwest to southeast warming pattern of temperatures observed in the state. Because of the large variability in temperature, the warm and cool seasons were defined from an agronomic perspective, primarily taking probabilities of freezing into account. Based on the climatological averages of Missouri, from 1983 to 2013, November through April registered average minimum temperatures below freezing, and was considered the cool season, while May through October's minimum average temperature were above freezing and constituted the warm season.

~~Missouri is characterized as a continental type of climate, marked by relatively strong seasonality. Furthermore, Missouri is subject to frequent changes in temperature, primarily due to its inland location and its lack of proximity to any large lakes. All of Missouri experiences below-freezing temperatures on a yearly-basis. For example, the majority of the state experiences, on average, 110 days with temperatures below freezing, while the Bootheel (i.e., southeast region) registers, on average, 70 days of below freezing day~~temperatures~~. This elaborates upon the typical northwest to southeast warming pattern of temperatures observed in the state. Because of the large variability in temperature, the warm and cool seasons were defined from an agronomic perspective, primarily taking probabilities of freezing into account. Based on the climatological averages of Missouri, from 1983 to 2013, November through April~~

~~registered average minimum temperatures below freezing, and was considered the cool season, while May through October's minimum average temperature were above freezing and constituted the warm season.~~

2.2 Rainfall data

In order for the results to be comparable across the domains of the three radars it was necessary to select days on which rain was observed widely across the state. Although measureable rainfall occurs on more than 100 days of the year in Missouri with only 50 days typically recording greater than 25.4 mm in 2014 had 46 days with measurable rainfall throughout the state. Furthermore, occurrence of rain was defined as the observation of an amount greater than 0.5 mm (equivalent to two rain gauge tips) in an hour. This amounted to a total of approximately 300 hours of rain across those 46 days. This represents a relatively standard year of rainfall for the state of Missouri. Furthermore, the days were chosen based on availability of data from the National Climate Data Center's (NCDC) Hierarchal Data Storage System (HDSS) for all three radars, in addition to error-free performance notes from each of the gauges used. The dates analyzed were split near evenly between warm (May – October) and cool (November – April), therefore encompassing an overall performance of each of the radars throughout the year with no preferential bias towards rain or snow. Additionally, days were distributed evenly during the summer between convective and stratiform events with a threshold of 38 dBZ (Gamache and Houze, 1982).

Terrestrial-based precipitation gauge data were collected from 15 separate weather stations within the Missouri Mesonet, established by the Commercial Agriculture Program of University Extension (Table 1). All precipitation data were aggregated in hourly intervals to match the temporal resolution of the gauges. Observed precipitation data were collected using Campbell Scientific TE525 tipping buckets located at each of the locations for the study (Table 1). The precipitation gauges have a 15.4 cm orifice which funnels to a fulcrum which registers 0.254 mm of rainfall per tip. The performance of each gauge is maximized between 0 and 50°C, for which each day of the study's temperature did not exceed. Accuracy in gauge measurements range between -1 to 1%, -3 to 0%, and -5 to 0% for precipitation up to 25.4 mm hr^{-1}, 25.4 to 50.8 mm hr^{-1}, and 50.8 to 76.2 mm hr^{-1}, respectively, which are, primarily, associated with local random errors and errors in tip-counting schemes (Kitchen and Blackall, 1992; Habib et al., 2001).

Each tipping bucket is located, approximately, 1 m above the ground in areas clear of buildings and properly maintained vegetation height to mitigate turbulence effects (Habib et al., 1999). Due to the well-maintained nature of the mesonet gauges, these errors were assumed negligible and, therefore, allowed for the gauges to be representative of the true rainfall rate. In spite of the non-homogeneous spacing of the gauges, unbiased statistics including the normalized mean bias and normalized standard error were utilized.

~~In order for the results to be comparable across the domains of the three radars it was necessary to select days on which rain was observed widely across the state. Although rainfall occurs on more than 100 days of the year in Missouri, in 2014 only 46 days had rain widespread enough for this study. Further to this, occurrence of rain was defined as the observation of an amount greater than 0.254 mm (equivalent to a single rain gauge tip) in an hour. This amounted to a total of approximately 300 hours of rain across those 46 days. This resultsrepresents in a relatively standard year of rainfall for the state of Missouri. Furthermore, the days were chosen based on availability of data from the National Climate Data Center's (NCDC) Hierarchal Data Storage System (HDSS) for all three radars, in addition to error-free performance notes from each of the gauges used. The dates analyzed were split near evenly between warm (May – October) and cool (November – April), therefore encompassing an overall performance of each of the radars throughout the year with no preferential bias towards rain or snow. Additionally, days were distributed evenly during the summer between convective and stratiform events with a threshold of 38 dBZ (Gamache and Houze, 1982).~~

~~Terrestrial-based (ground-truthed) precipitation gauge data were collected from 15 separate weather stations within the Missouri Mesonet, established by the Commercial Agriculture Program of University~~

Formatted: Indent: First line: 0"

Formatted: Not Highlight

~~Extension (Table 1). All precipitation data were aggregated in hourly intervals to match the temporal resolution of the ground-truthed gauges. Forty-six out of 365 days for the year of 2014 were analyzed based on precipitation being registered across the entire study domain (Figure 1). Of these 46 days, approximately 300 out of 1,104 hours of precipitation occurred such that the tipping buckets recorded more than one tip (i.e., greater than 0.254 mm) for each location. This results in a relatively standard year of rainfall for the state of Missouri. Furthermore, the days were chosen based on availability of data from the National Climate Data Center's (NCDC) Hierarchal Data Storage System (HDSS) for all three radars, in addition to error-free performance notes from each of the gauges used. The dates analyzed were split near evenly between warm (May – October) and cool (November – April), therefore encompassing an overall performance of each of the radars throughout the year with no preferential bias towards rain or snow. Additionally, days were distributed evenly during the summer between convective and stratiform events with a threshold of 38 dBZ (Gamache and Houze, 1982).~~

~~Observed precipitation data were collected using Campbell Scientific TE525 tipping buckets located at each of the locations for the study (Table 1). The precipitation gauges have a 15.4 cm orifice which funnels to a fulcrum which registers 0.01 mm of rainfall per tip. The performance of each gauge is maximized between 0 and 50°C, for which each day of the study's temperature did not exceed. Accuracy in gauge measurements range between -1 to 1%, -3 to 0%, and -5 to 0% for precipitation up to 25.4 mm hr^{-1}, 25.4 to 50.8 mm hr^{-1}, and 50.8 to 76.2 mm hr^{-1}, respectively, which are, primarily, associated with local random errors and errors in tip-counting schemes (Kitchen and Blackall, 1992; Habib et al., 2001).~~

~~Each tipping bucket is located, approximately, 1 m above the ground in areas clear of buildings and properly maintained vegetation height to mitigate turbulence effects (Habib et al., 1999). Due to the well-maintained nature of the mesonet gauges, these errors were assumed negligible and, therefore, allowed for the gauges to be representative of the true rainfall rate. In spite of the non-homogeneous spacing of the gauges, unbiased statistics including the normalized mean bias and normalized standard error were utilized.~~

Formatted: Highlight

Formatted: Highlight

~~2.3 Radar data and radar-rainfall algorithms~~

~~Next Generation Radar (NEXRAD) level-II data were retrieved from the NCDC's HDSS. Files were analyzed~~ processed ~~using the Weather Decision Support System – Integrated Information (WDSS-II) program (Lakshmanan et al., 2007a) to assess reflectivity (Z) in addition to dual-polarized radar variables including differential reflectivity (ZDR) and specific differential phase shift (KDP). Many different quality control techniques are available (e.g., Lakshmanan et al., 2007b, 2010, 2014) and implemented upon the radar data with WDSS-II. Three other variables were also generated based on a KDP-based smoothing field (Ryzhkov et al., 2003) for reflectivity, differential reflectivity, and specific differential phase: DSMZ, DZDR, and DKDP, respectively. These were implemented to determine whether the additional KDP-smoothing fields tend to over- or underestimate QPE's (Simpson et al., 2016). A rain rate echo classification variable (RREC) was also computed, which chooses whether an R(Z), R(KDP), R(Z,ZDR), or R(ZDR, KDP) algorithm is implemented in estimating rain rates based on the radar fields of Z, ZDR, and KDP (Kessinger et al., 2003) to determine whether a multi-parameter algorithm is superior to a single algorithm.~~

~~All seven variables (Z, ZDR, KDP, DSMZ, DZDR, DKDP, and RREC) were converted from their native polar grid to 256 x 256 1 km Cartesian grids, where the lowest radar elevation scans (0.5°) were used to mitigate uncalculated effects from evaporation and wind drift. An average of 5 minute scans were used for each of the variables, which were aggregated to hourly totals to be compared to the hourly tipping-bucket accumulations. In spite of previous reports suggesting 5 minute to hourly aggregates can have significant effects on QPE (e.g., Fabry et al. 1994), Shucksmith et al. (2011) present evidence that accumulation overestimation did not exceed 26% for a pixel size of 1 km.~~

~~The latitude and longitude of each of the 15 gauges were matched with the radar pixel that corresponds to the Cartesian grid value of the seven radar variables which were then implemented in rain rate~~

Formatted: Indent: Left: 0", First line: 0"

Formatted: Indent: First line: 0"

~~calculations. These rain-rate calculations were calculated using the equations presented by Ryzhkov et al. (2005) (Table 2), which were gathered from multiple studies using disdrometers to derive a relationship between reflectivity, differential reflectivity, and specific differential phase (Bringi and Chandrasekar, 2001; Brandes et al., 2002; Illingworth and Blackman, 2002; Ryzhkov et al., 2003). Standard R(Z) algorithms were also included to test whether the addition of dual-polarized technology improves QPE's.~~

~~With the use of both Z, ZDR, KDP, and DSMZ, DZDR, and DKDP fields produced by WDSS-II, the number of algorithms tested was 55. This includes the three standard single-polarized algorithms (stratiform, convective, and tropical) which were calculated using reflectivity R(Z), and then calculated as R(DSMZ), while algorithms 1-6 (R(KDP)) were also calculated as R(DKDP). Algorithms 7-11 (R(Z, ZDR)) were additionally calculated as R(Z, DZDR), R(DSMZ, ZDR), and R(DSMZ, DZDR), while the same four combinations of non- and KDP-smoothed fields were applied to the R(KDP, ZDR) algorithms (12-15).~~

~~2.4 Statistical analyses~~

~~To test the performance of each algorithm, several statistical analyses were calculated. The average difference (Bias) was calculated as~~

$$Bias = \frac{\sum (R_i - G_i)}{N} \qquad (1)$$

~~where R_i is each hourly aggregated radar estimated rainfall amount calculated from one of the 55 algorithms, G_i is the hourly aggregated gauge (observed) measurement, and N is the total number of observations which, for this study, was 1,104 hours. A second statistical parameter, the normalized mean bias (NMB), was calculated as~~

$$NMB = \frac{1}{N}\frac{\sum (R_i - G_i)}{\sum G_i} \qquad (2)$$

Formatted: Highlight

~~The normalized mean bias is included in the analyses due to the fact that overestimations (i.e., radar estimates larger than gauge measurements) and underestimations (i.e., radar estimates smaller than gauge measurements) are treated proportionately. This is directly analogous to choosing the mean absolute error (MAE) opposed to the standard deviation as the MAE does not penalize smaller or larger errors, obscuring the overall results (Chai and Draxler, 2014). Bias measurements (Bias and NMB) were calculated to determine whether radar derived rain rates were over- or under-estimated in comparison to the gauges. However, to calculate the overall magnitude of error associated with the performance of the radars, the absolute values of (1) and (2) were performed to yield the mean absolute error (MAE), and normalized standard error (NSE), respectively.~~

~~Several other meteorological parameters were calculated, including probability of detection (PoD) which was calculated as~~

$$PoD = \frac{\sum |R_i \bullet G_i > 0 \,\&\, R_i > 0|}{\sum |G_i|} \qquad (3)$$

~~where the bullet (•) indicates "if", to determine how accurate the radars were at correctly detecting precipitation. The probability of detection values range between 0.0 (radar did not detect any precipitation correctly) and 1.0 (radar detected the occurrence of all precipitation 100% correctly). The probability of false detection takes into account the amount of precipitation the radars incorrectly estimated when the gauges recorded zero values, and was calculated as~~

$$PoFD = \frac{\sum R_i \bullet (G_i = 0 \,\&\, R_i > 0)}{\sum G_i} \qquad (4)$$

~~Quantitative measures including the missed precipitation amount (MPA) and the false precipitation amount (FPA) were defined such that~~

$$MPA = \sum R_i \bullet (G_i > 0 \,\&\, R_i = 0) \qquad (5)$$

$$\cancel{FPA = \sum R_i \bullet (G_i = 0 \,\&\, R_i > 0)} \qquad \cancel{(6)}$$

~~which analyzes the total amount of precipitation due to misses and false alarms. The total precipitation error was also recorded to assess the overall error from each radar.~~

~~3 Results and discussion~~

~~3.1 Overall algorithm performance~~

~~To test the overall performance of each radar, it was necessary to determine the overall best algorithm for each statistical measure. The best algorithm from each grouping of equations was determined to have the lowest normalized standard error (NSE), indicating the best performance relative to the gauge-recorded precipitation amount (Ryzhkov et al., 2005). This reduces the impact of bias inherent within the dataset between warm/cool season, stratiform/convective events, and allows for statistical measurements in spite of the (typical) non-Gaussian behavior of precipitation (Kleiber et al., 2012; Alaya et al., 2017).~~

~~From the results obtained, the three R(Z) three R(DSMZ) displayed a particular bias in favor of the R(Z)-Convective algorithm for all three radars with R(Z)-Stratiform displaying similar performance (Figure 2a). This could be due, at least in part, to the near-equal stratiform and convective precipitation regimes throughout 2014. Although errors generally increased as range increased for KEAX and KLSX, the results were nebulous for KSGF. The lowest NSE values were, typically, closest to each of the radars (between 0.4 and 0.8), with the notable exception of the closest gauge to KSGF. In general, the RREC performed worst, potentially due to the algorithm's ability to correctly assess the hydrometeors present~~

Formatted: Indent: Left: 0", First line: 0"

Formatted: Indent: Left: 0"

Formatted: Indent: First line: 0"

~~(Cifelli et al., 2011; Yang et al. 2016). Additionally, the poor performance by the R(DSMZ)-Tropical equation is due to the lack of tropical precipitation within Central Missouri. Overall, the KDP-smoothed reflectivity fields (DSMZ) performed worse than their counter-parts, resulting in over-prediction of precipitation and, thus, larger errors (Simpson et al., 2016). Errors did not exceed 2.4 for any of these algorithms.~~

~~However, the performance of the KDP-smoothed KDP field (DKDP) performed better than the original specific differential phase shift field (Figure 2b). For nearly all gauges for each of the 3 radars, R(DKDP)4 performed the best, with NSE values ranging from 1.4 to 4.1. The range of NSE values were largest at KEAX, while the spread was relatively small for KLSX and KSGF. In spite of this, the overall spread of the performance of the 12 KDP algorithms varied greatly (average of 2 NSE units), exhibiting the sensitivity of KDP estimates on QPE (Ryzhkov et al., 2005; Cunha et al., 2013). In general, the NSSL-derived R(KDP) equations (i.e., equations 4-6) outperformed those from Bringi and Chandrasekar (2001, equation 1), Brandes et al. (2002, equation 2), and Illingworth and Blackman (2002, equation 3). Regardless, the magnitudes were all, approximately, more than 1 NSE unit than the performance of the R(Z) algorithms.~~

~~The algorithms with the lowest NSE values were equations 7-11. For example, the overall lowest NSE was at a distance of 130 km from KEAX (0.3), with no locations exceeding NSE values of 2.0 (Figure 2c). The large values at the closest location for KSGF (85 km, 1.3 – 1.9 NSE units), and the fifth closest gauge to KLSX (135 km, 1.3 – 1.8 NSE units), Cook Station, were similar to the R(Z) and R(DSMZ) results, indicating potential issues with reflectivity measurements. Additionally, these locations were the closest in performance to the R(KDP) and R(DKDP) NSE values. Observations from this gauge (Cook Station) indicated hail occurred during the evening of 01 August, for which KDP estimates would be more ideal than Z for QPE (Ryzhkov et al. 2005; Kumjian 2013a; Cunha et al. 2015). In spite of this, the overall spread in performance of the R(Z,ZDR) equations were less than the R(KDP) equations,~~

~~demonstrating the robust performance of R(Z,ZDR) for QPE (Wang and Chandrasekar 2010; Seo et al., 2015).~~

~~The R(ZDR,KDP) algorithms performed the worst, overall (Figure 2d). In spite of the differential reflectivity being implemented, the overall NSE values increased in magnitude, exceeding 6 units for the second gauge analyzed by KEAX. Algorithms containing DKDP measurements performed better than simply KDP, demonstrating that even with the scaling behavior of ZDR, DKDP is superior to KDP estimates. This provides a potential solution to the noisy-ness that tends to be exhibited in the KDP field (Ruzanski and Chandrasekar 2012).~~

~~Due to the overall NSE values obtained, for the remainder of the analyses, equation 11 (i.e., R(Z,ZDR)5) and equation 13 (i.e., R(ZDR,KDP)2) will be utilized as the best and worst algorithms, respectively. Equations containing DZDR were not included in the following discussion due to the very large QPE errors for each radar.~~

Formatted: Indent: Left: 0"

~~3.2 KEAX~~

Formatted: Indent: First line: 0"

~~The overall bias showed that there was a positive bias, peaking near 5.5 mm hr~~$^{-1}$ ~~at the second gauge for KEAX, approximately 115 km from the radar for both the best and worst performing algorithms (Figure 3). This corresponds well with the spike in falsely detected precipitation recorded, which is canceled by the maximum in missed precipitation at the second distance of, approximately, 150 km. The overall worst algorithm, equation 13, an R(ZDR,KDP) relationship, revealed a decreasing trend in bias as the distance from the radar increased. For example, a bias of 4 mm hr~~$^{-1}$ ~~was observed at a distance of 75 km from the radar, whereas the bias reduced to 3 mm hr~~$^{-1}$ ~~at distances near 175 km. This could be due, at least in part, to the algorithm's utilization of KDP which performs poorly in frozen (especially light) precipitation (Zrnic and Ryzhkov, 1996; Kumjian 2013a), causing the overestimation. Conversely, the algorithm with the lowest bias was an R(Z,ZDR) algorithm (equation 11). There was a maximum in the bias calculations~~

~~while utilizing equation 11 near 120 km, similar to equation 13, however, there was a more pronounced minimum in the data near 150 km. Furthermore, it appears the data oscillates around a bias value of 0 mm hr^{-1} when using equation 13. This could be due to ZDR's capability to respond to precipitation shape (Kumjian 2013a), which helps to scale the reflectivity portion of the rainfall estimation algorithm to a more accurate value (Seo et al., 2015). In general, the cool season displayed a larger magnitude of error in terms of bias for both algorithms.~~

~~The normalized mean bias (NMB) reveals the same trend in values for bias but with an overall decrease in magnitude. It is important to note, however, that the algorithms that tend to perform the worst (e.g., algorithms containing KDP) result in anomalous range responses which would be due, at least in part, to a stronger response to precipitation type. This indicates that observations above the melting layer are dominant for which QPE's tend not to be calculated (Cifelli et al., 2011; Seo et al., 2015) but are important for regions devoid of adequate radar coverage (Ryzhkov et al., 2003; Simpson et al., 2016).~~

~~The absolute bias and normalized standard error (NSE) shows the same maxima in the data at the second gauge (Brunswick) that was present in the bias data (6.2 mm hr^{-1} and 5.6, respectively) . However, a second maxima is located at the fifth gauge at, approximately, 150 km (Linneus) with values of 5.9 mm hr^{-1} and 4.0, respectively. Bright-band issues are detected due, at least in part, to the increased missed precipitation amount (240 mm) at this particular distance for the R(ZDR,KDP) equation (i.e., worst performing algorithm). There was also a pronounced minimum in the absolute bias and NSE results at the fourth gauge for equations 11 and 13, 4.0 mm hr^{-1} and 0.8 mm hr^{-1}, and 2.8 and 0.8, respectively, potentially indicating an idealized range of QPE for KEAX. Furthermore, the historical records at this particular gauge showed less issues (e.g., clogging) than any of the others analyzed by the KEAX radar. This highlights the importance of choosing ground-truth data, in particular tipping buckets which are prone to numerous errors (Ciach and Krajewski, 1999b).The largest contributions to the NSE and NMB were due to the warm season.~~

~~The probability of detection (PoD) results indicate a large difference in algorithm choice for correctly detecting precipitation. The low PoD at, approximately 150 km, indicates overshooting of the beam. This is further evidenced by the MPA results, as about 225 mm of precipitation was missed by the radar at 150 km, whereas only 100 mm of precipitation was missed by the radar at the second gauge at 120 km. Although equation 11, an R(Z,ZDR) algorithm was superior in terms of the bias, the same algorithm with a KDP-smoothed reflectivity value, R(DSMZ,ZDR) revealed the overall least amount of falsely missed precipitation (by 10 mm). However, the summation of the amount of precipitation falsely detected (PoFD) by KEAX showed a larger source of error than the MPA in terms of magnitude. For example, at the second (fifth) gauge, only 100 (225) mm of precipitation was missed by the radar, but over 700 (725) mm of precipitation was incorrectly estimated by the radar.~~

~~Correlation coefficient (CC) values for any of the 9 stations analyzed by KEAX ranges from 0.02 (Linneus, 151 km) to 0.93 for the cool season (St. Joseph, 115 km). The lowest~~ ~~R^2~~ ~~were due to a combination of false alarms and misses. For example, the CC for the warm seasons at Sanborn (170 km) and Jefferson Farm (173 km) were 0.22 and 0.24, respectively, whereas when the instances of false alarms and misses were removed, increased to 0.48 and 0.52. Few locations (Brunswick, 114 km and Versailles, 129 km) saw little improvement in the CC values when only hits were analyzed (less than 0.1 increase), indicating the mean absolute error (in terms of hits) contributed the largest portion of error.~~

Formatted: Indent: Left: 0"

~~3.3 KLSX~~

~~Unlike the KEAX data, the gauges used for analyses for the KLSX radar span between 90 – 150 km. Furthermore, 5 out of the 8 gauges were located within 10 km of range from one another, near 140 km from the radar, limiting the data available for analyses between 100 and 140 km (Figure 5).~~

Formatted: Indent: First line: 0"

~~The bias and NMB both show a relatively modest peak in values near the second gauge of 5 mm , which decreases to approximately 3.6 mm at the third gauge, 120 km from the radar. The worst performing~~

~~algorithm, equation 13, was the same R(ZDR,KDP) relation as the worst KEAX bias and NMB data. Additionally, the overall trend of decreasing bias and NMB as distance from the radar increases was noted, presumably due to overshooting effects similar to the KEAX data. Furthermore, the overall non-biased results from the R(Z,ZDR) equation demonstrates its robust capabilities in QPE, in spite of its sensitivity to calibration (Zrnic et al., 2005; Bechini et al., 2008).~~

~~The double maxima in the absolute bias graph are present as with the KEAX data, but are not as pronounced. For example, the absolute bias at 95 km and 140 km from KLSX were 5.9 mm and 1.1 mm , and 4.9 mm and 1.4 mm for equations 13 and 11, respectively. Additionally, the overall minima in the absolute bias for both KEAX and KLSX are at, approximately, 125 km from the radar (3.9 mm hr~~$^{-1}$ ~~and 1.0 mm hr~~$^{-1}$~~, respectively, for equations 13 and 11). The relative distance from the radars are the same, where the two maxima for KEAX were at 115 and 150 km, while the maxima were at, approximately, 100 and 140 km for KLSX. The overall best and worst performing algorithms at KLSX for the absolute bias and NSE were equations 11 and 13, the R(Z,ZDR) and R(ZDR,KDP) algorithms, respectively.~~

~~The magnitude of error in terms of absolute bias, normalized mean bias, and normalized standard error, all showed a decreasing pattern as distance from KLSX increased. This was due, primarily, from a maximum in the false precipitation amount at 95 km from the radar. Historical notes at this location indicate frequent clogging of the rain gauge, either due to bugs or leaves. From a particular series of events spanning from 01 to 04 April and 01 to 03 August, 2014, over 130 mm of precipitation occurred during each period which was not captured by the gauge, resulting in a large amount of overall error. These results indicate the important of dual gauges in the same vicinity (Krajewski et al. 1998; Ciach and Krajewski 1999). Interestingly, the cool season displayed a larger NSE (5 % for R(ZDR,KDP)) potentially due to the very low probability of detection (0.2) at this range of 118 km.~~

~~One of the main differences between the KLSX and KEAX data was the decreased probability of detection at 120 km for KLSX, while there was an increased probability of detection for KEAX. In general, the PoD values were worse for KLSX when compared to KEAX. For example, equation 11 had~~

~~no PoD values below 0.90, whereas no PoD values exceeded 0.84 for KLSX. There was also a slight trend of increasing PoD values as distance from the St. Louis radar increased and, at one point near 140 km, the best algorithm, R(DSMZ) convective and the worst algorithm, KDP1, were not significantly different ($p < 0.10$). Additionally, the maxima in the PoD while utilizing KDP1 corresponds to a minima in the R(DSMZ) detection percentage, which is well correlated by the similarly valued MPA results.~~

~~The missed precipitation amount (MPA) displayed the cool season contributed the most, whereas the warm season contributed the most amount of false precipitation amount. The R(Z,ZDR) equation only registered, on average, 25 mm of MPA and 160 mm of FPA, whereas the R(ZDR,KDP) equation was very dependent upon range. For example, the FPA from R(ZDR,KDP) decreased as range increased from the radar from a maximum of, approximately, 850 mm to 620 mm. However, the fifth-furthest gauge (137 km from KLSX) displayed a sharp increase in the MPA for both cool seasons (above 100 mm).~~

~~3.4 KSGF~~

~~In spite that the KLSX and KEAX data strongly suggests false precipitation errors near 100 km in addition to bright-banding near 150 km from the radars, the KSGF results reveal an overall smooth decrease (increase) of error with range (Figure 7) for R(ZDR,KDP) and R(Z,ZDR), accordingly. One of the main reasons for this could be due to the fact that only 5 gauges were analyzed from KSGF (the fewest of the 3 radars analyzed), smoothing the overall trend lines.~~

~~The bias remained relatively constant near -0.3 mm for R(Z,ZDR), whereas the bias exhibited a sharp decrease from 4 mm to 2.7 mm over a distance of, approximately, 100 km. In general, the cool season~~

Formatted: Indent: First line: 0"

~~displayed the lower of bias magnitudes when compared to the warm season, similar to the KEAX results. This may be due, at least in part, to the low PoFD values for the warm season close to the KSGF radar.~~

~~Similar to the bias, the absolute bias for R(Z,ZDR) was constant at all ranges (near 1 mm) whereas the R(ZDR,KDP) equation decreased from 5.2 mm to 3.8 mm. This is potentially due to the low cool season PoD values (below 0.6), while the warm season R(ZDR,KDP) values (near 0.8) remained constant. A larger contribution from more correctly detected precipitation in addition to the decreasing trends in the NMB and NSE would result in a lower absolute bias.~~

~~The closest location (90 km) typically displayed the largest errors for the R(ZDR,KDP) equation, and then decreased in error magnitude as range increased. In spite of this, the PoFD results indicate both algorithms increased in PoFD values as range increased, with the warm season typically dominating, particularly due to the large convective clouds dominate in the warm season. False detection values as low as 0.01 for the cool season while utilizing R(Z,ZDR) were observed at distances near 100 km and 140 km from the radar.~~

~~Normalized standard error values increased from 0.7 % at a distance of 105 km to 1.8 % at a distance of 185 km for R(Z,ZDR). Large NSE values for the warm season (7.5 %) were calculated for R(ZDR,KDP) which decreased to 3.8 % at 185 km from the radar. Furthermore, this was the only instance when the warm season was less than the cool season in terms of NSE. Otherwise, the overall NSE decreased from 5 % to 3.9 % for R(ZDR,KDP). The NMB followed a similar trend for the KDP-containing algorithm, with a noticeable exception at the second gauge (105 km from KSGF), where the overall NSE was closer to the warm than cool season. This is due to the low PoFD values at this location, in addition to a smaller difference between the two algorithm's FPA measurements.~~

~~The MPA results, unlike for KEAX and KLSX, displayed a larger range of performance between seasons. However, the warm season still exhibited the overall best performance in terms of MPA, yet contributed the most to the FPA for both R(Z,ZDR) and R(ZDR,KDP). In spite of the MPA typically increasing as~~

~~range increased, the FPA was more nebulous. For example, the second gauge (105 km from KSGF) had the overall lowest NSE (0.8 %), MPA (15 mm), and FPA (95 mm) for R(Z,ZDR). The third-furthest location (142 km) resulted in slightly larger errors, overall, while the fourth-furthest location had errors similar to the second gauge (105 km). Then, at the furthest tipping bucket location (185 km), NSE values increased, whereas FPA and MPA decreased. Therefore, the furthest location's errors are due, primarily, from discrepancies between precipitation magnitude between the gauge and radar.~~

~~Excluding Versailles (142 km from KSGF), the cool season exhibited larger~~ R^2 ~~values in comparison to the cool season (Figure 8). Furthermore, CC values exceeded 0.9 when false alarms and misses were excluded from Mt. Grove (101 km) and was 0.84 when included. Otherwise, the other four stations analyzed by the Springfield radar displayed many counts of false alarms and misses, leading to low~~ R^2 ~~values.~~

~~Due to the relatively large ranges from the Springfield (KSGF) radar, most of the correlation coefficient values were low in comparison to either KLSX or KEAX. For the warm (cool) season without false alarms and misses,~~ R^2 ~~values ranged from 0.44 (0.38) and 0.34 (0.36) for KLSX and KSGF, respectively, at Cook Station (119 and 185 km). Similarly, the CC values ranged from 0.61 (0.71) to 0.42 (0.56) at Green Ridge (76 and 154 km) for KEAX and KSGF, accordingly.~~

Formatted: Indent: First line: 0"

~~4 Conclusions~~

Formatted: Indent: Left: 0", First line: 0"

~~Dual-polarization technology was implemented to the National Weather Service Next Generation Radar network in the Spring of 2012 to, primarily, improve quantitative precipitation estimation and hydrometeor classification. The current study observed over 1,100 hours of precipitation data with three~~

Formatted: Indent: First line: 0"

~~separate radars in Missouri using 55 algorithms including the three conventional R(Z) radar rain-rate estimation algorithms (stratiform, convective, and tropical) along with a myriad of R(KDP), R(Z,ZDR), and R(ZDR,KDP) algorithms which can be found in Ryzhkov et al. (2005). Additionally, a KDP-smoothing field of reflectivity, differential reflectivity, and the specific differential phase shift (DSMZ, DZDR, and DKDP, respectively) were measured and used for analyses. Unlike previous studies, the current work emphasizes the amount of precipitation correctly and incorrectly estimated by the radar in comparison to the terrestrial based precipitation gauges through measurements of the missed and false precipitation amount.~~

~~For all three radars, Kansas City, St. Louis, and Springfield, MO (KEAX, KLSX, and KSGF, respectively), the majority of precipitation error (over 60%) was contributed by the amount of precipitation falsely detection by the radar (up to 725 mm), while 20% was due to the radar missing the precipitation (up to 225 mm) for KEAX. Similar magnitudes of error were reported for KLSX and KSGF, with an overall error in precipitation for each radar ranging between 250 mm for the best performing of the 55 algorithms, equation 11 (an R(Z,ZDR) algorithm), and up to 2000 mm for the worst performing algorithms, R(ZDR,KDP) equation 13. The R(Z,ZDR) equation (an NSSL algorithm) was determined to be the most robust due to it registering the lowest NSE.~~

~~The data was divided into summer (May – October) and winter (November – April) months resulting in 652 hours for summer, and 452 hours for winter (59 and 41% of the entire data, respectively). Despite the winter data contributing less than the summertime data, it accounted for 20% of the overall MPA, and 40% to the overall PoFD. The~~ $\cancel{R^2}$ ~~values were less during the winter in comparison to the warm season primarily due to the smaller magnitude of precipitation that occurred. Furthermore, CC values increased by as much as 0.4 when instances of hits and misses were removed from the analyses, resulting in the warm season to outperform the cool season CC values at particularly short ranges from the radar.~~

~~These results aid in our understanding in the possibilities for hydrometeorological studies. Nearly 50% of the 1,100 hours analyzed for the study consisted of either falsely estimated precipitation by the radar, or~~

~~missed by the radar. Furthermore, these errors accumulate between 500 to 2,000 mm of precipitation depending on the algorithms chosen. Although the overall performance increased when false alarms and misses were removed, correlation coefficient values still, typically, remained below 0.50 at ranges beyond 130 km.~~

~~Furthermore, results demonstrate the issues with analyzing QPE from a single gauge, explaining why the Community Collaborative Rain, Hail, and Snow Network (Kelsch 1998; Cifelli et al., 2005; Reges et al., 2016) tends to be more utilized since results have shown that measurements or quality controlled techniques made by CoCoRaHS are significantly more accurate than rain gauges (Simpson et al., 2017), especially for convective events (Moon et al. 2009).~~

~~**Author Contribution.** N. Fox designed the experiment and provided feedback while M. Simpson carried out the calculations and wrote the manuscript.~~

~~**Acknowledgements.** This material is based upon work supported by the National Science Foundation under Award Number IIA-1355406. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.~~

Formatted: Indent: Left: 0"

~~References~~

~~Alaya, M.A., Ourda, T.B.M.J., Chebana, F.: Non-Gaussian spatiotemporal simulation of multisite precipitation: Downscaling framework. Climate Dynamics, 2017. doi: https://doi.org/10.1007/s00382-017-3578-0.~~

~~Anagnostou, M.N., Anagnostou, E.N., Vulpiani, G., Montopoli, M., Marzano, F.S., Vivekanandan, J.: Evaluation of X-band polarimetric-radar estimates of drop-size distributions from coincident S-band~~

Formatted: Indent: Left: 0", First line: 0"

~~polarimetric estimated and measured raindrop spectra. IEEE Transactions on Geoscience and Remote Sensing, 46, 3067-3075, 2008.~~

~~Bechini, R., Baldini, L., Cremonini, R., Gorgucci, E.: Differential reflectivity calibration for operational radars. Journal of Atmospheric and Oceanic Technology, 25, 1542-1555, 2009.~~

~~Berne, A. and Uijlenhoet, R.: A stochastic model of range profiles of raindrop size distributions: application to radar attenuation correction, Geophysical Research Letters, 32, 2005, doi: https://doi.org/10.1029/2004GL021899.~~

~~Berne, A. and Krajewski, W.F.: Radar for hydrology: Unfulfilled promise or unrecognized potential? Advances in Water Resources, 51, 357-366, 2013.~~

~~Bringi, V.N. and Chandrasekar, V.: Polarimetric Doppler weather radar, principles and applications. Cambridge University Press: Cambridge, UK, 636, 2001.~~

~~Brandes, E.A., Zhang, G., Vivekanandan, J.: Experiments in rainfall estimation with a polarimetric radar in a subtropical environment, Journal of Applied Meteorology, 41, 674–685, 2002.~~

~~Brandes, E.A., Zhang, G., Vivekanandan, J.: Drop size distribution retrieval with polarimetric radar: model and application, Journal of Applied Meteorology, 43, 461-475, 2004.~~

~~Chai, T., Draxler, R.R.: Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature, Geoscientific Model Development, 7, 1247-1250, 2014.~~

~~Ciach, G.J., Krajewski, W.F.: On the estimation of radar rainfall error variance. Advances in Water Resources, 22, 585-595, 1999a.~~

~~Ciach, G.J. and Krajewski, W.F.: Radar-raingage comparisons under observational uncertainties. Journal of Applied Meteorology, 38, 1519-1525, 1999b.~~

~~Ciach, G.J.: Local random errors in tipping-bucket rain gauge measurements. Journal of Atmospheric and Oceanic Technology, 20, 752-759, 2002.~~

~~Cifelli, R., Doesken, N., Kennedy, P., Carey, L.S., Rutledge, S.A., Gimmestad, C., Depue, T.: The community collaborative rain, hail, and snow network: Informal education for scientists and citizens. Bulletin of the American Meteorological Society, 86, 1069-1077, 2005.~~

~~Cunha, L.K., Smith, J.A., Baeck, M.L., Krajewski, W.F.: An early performance of the NEXRAD dual-polarization radar rainfall estimates for urban flood applications. Weather and Forecasting, 28, 1478-1497, 2013.~~

~~Cunha, L.K., Smith, J.A., Krajewski, W.F., Baeck, M.L., Seo, B.: NEXRAD NWS polarimetric precipitation product evaluation for IFloods. Journal of Hydrometeorology, 16, 1676-1699, 2015.~~

~~Delrieu, G., Andrieu, H., Creutin, J.D.: Quantification of path-integrated attenuation for X- and C-band weather radar systems operating in Mediterranean heavy rainfall. Journal of Applied Meteorology, 39, 840-850, 2000.~~

~~Fabry, F., Bellon, A., Duncan, M.R., Austin, G.L.: High resolution rainfall measurements by radar for very small basins: the sampling problem reexamined. Journal of Hydrology, 161, 415-428, 1994.~~

~~Fairman, J.G., Schultz, D.M., Kirschbaum, D.J., Gray, S.L., Barrett, A.I.: A radar-based rainfall climatology of Great Britain and Ireland. Weather, 70, 153-158, 2012. doi: https://doi.org/10.1002/wea.2486.~~

~~Gamache, J.F. and Houze, R.A.: Mesoscale air motions associated with a tropical squall line. Monthly Weather Review, 110, 118–135, 1982.~~

~~Giangrande, S.E. and Ryzhkov, A.V.: Estimation of rainfall based on the results of polarimetric echo classification. Journal of Applied Meteorology, 47, 2445-2460, 2008.~~

~~Gorgucci, E., Scarchilli, G., Chandrasekar, V.: Calibration of radars using polarimetric techniques. IEEE Transactions in Geoscience and Remote Sensing, 30, 853-858, 1992.Gorgucci, E., Scarschilli, G., Chandrasekar, V., Bringi, V.N.: Measurement of mean raindrop shape from polarimetric radar observations. Journal of the Atmospheric Sciences, 57, 3406-3413, 2000.~~

~~Gorgucci, E., Baldini, L., Chandrasekar, V.: What is the shape of a raindrop? An answer from radar measurements. Journal of the Atmospheric Sciences, 63, 3033-3044, 2006.~~

~~Goudenhoofdt, E., Delobbe, L.: Long-term evaluation of radar QPE using VPR correction and radar-gauge merging. International Association of Hydrological Sciences Publications, 351, 249-254, 2012.~~

~~Goudenhoofdt, E., Delobbe, L.: Generation and verification of rainfall estimates from 10-yr volumetric weather radar measurements. Journal of Hydrometeorology, 133, 1191-1204, 2016.~~

~~Gourley, J.J., Giangrande, S.E., Hong, Y., Flamig, Z., Schuur, T., Vrugt, J.: Impacts of polarimetric radar observations on hydrologic simulation. Journal of Hydrometeorology, 11, 781-796, 2010.~~

~~Habib, E., Krajewski, W.F., Nespor, V., Kruger, A.: Numerical simulation studies of rain gauge data correction due to wind effect. Journal of Geophysical Research, 104, 723–734, 1999.~~

~~Habib, E., Krajewski, W.F., Kruger, A.: Sampling errors of tipping-bucket rain gauge measurements. Journal of Hydrological Engineering, 6, 159–166, 2001.~~

~~Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D., New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. Journal of Geophysical Research, 113, 2008. doi: https://doi.org/10.1029/2008JD010201~~

~~Holleman, I., Huuskonen, A., Gill, R., Tabary, P.: Operational monitoring of radar differential reflectivity using the sun. Journal of Atmospheric and Oceanic Technology, 27, 881-887, 2010.~~

~~Hubbert, J.C.: Differential reflectivity calibration and antenna temperature. Journal of Atmospheric and Oceanic Technology, 34, 1885-1906, 2017.~~

~~Illingworth, A., Blackman, T.A.: The need to represent raindrop size spectra as normalized gamma distributions for the interpretation of polarization radar observations. Journal of Applied Meteorology, 41, 286-297, 2002.~~

~~Kelsch, M.: The Fort Collins flash flood: Exceptional rainfall and urban runoff. Preprints, 19th Conference on severe local storms, Minneapolis, MN, American Meteorological Society, 404-407, 1998.~~

~~Kitchen, M. and Blackall, M.: Representativeness errors in comparisons between radar and gauge measurements of rainfall. Journal of Hydrology, 134, 13–33, 1992.~~

~~Kleiber, W., Katz, R.W., Rajagopalan, B.: Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. Water Resources Research, 48, 2012. doi: https://doi.org/10.1029/2011WR011105.Kessinger, C., Ellis, S., Van Andel, J.: The radar echo classifier: a fuzzy logic algorithm for theWSR-88D. 19th Conf. on Inter. Inf. Proc. Sys. (IIPS) for Meteor., Ocean., and Hydr., Amer. Meteor. Soc., Long Beach, CA, 2003.~~

~~Kitchen, M. and Jackson, P.M.: Weather radar performance at long range – simulated and observed. Journal of Applied Meteorology, 32, 975-985, 1993.~~

~~Krajewski, W.F., Kruger, A., Nespor, V.: Experimental and numerical studies of small-scale rainfall measurements and variability. Water Science and Technology, 37, 131-138.~~

~~Kumjian, M.R.: Principles and applications of dual-poarization weather radar. Part 1: Description of the polarimetric radar variables. Journal of Operational Meteorology, 1, 226-242, 2013a.~~

~~Kumjian, M.R.: Principles and applications of dual-poarization weather radar. Part 2: Warm and cold season applications. Journal of Operational Meteorology, 1, 243-264, 2013b.~~

~~Kumjian, M.R.: Principles and applications of dual-poarization weather radar. Part 3: Artifacts. Journal of Operational Meteorology, 1, 265-274, 2013c.~~

~~Lakshmanan, V., Smith, T., Stumpf, G., Hondl, K.: The warning decision support system—integrated information. Weather and Forecasting, 22, 596–612, 2007a.~~

~~Lakshmanan, V., Fritz, A., Smith, T., Hondl, K., Stumpf, G.: An automated technique to quality control radar reflectivity data. Journal of Applied Meteorology and Climatology, 46, 288-305, 2007b.~~

~~Lakshmanan, V., Zhang, J., Howard, K.: A technique to censor biological echoes in radar reflectivity data. Journal of Applied Meteorology and Climatology, 49, 453-462, 2010.~~

~~Lakshmanan, V., Karstens, C., Krause, J., Tang, L,: Quality control of weather radar data using polarimetric variables. Journal of Atmospheric and Oceanic Technology, 31, 1234-1249, 2014.~~

~~Moon, J.T., Guinan, P.E., Snider, D.J., Lupo, A.R.: CoCoRaHS in Missouri: Four years later, the importance of observations. Transactions of the Missouri Academy of Science, 43, 7-18, 2009.~~

~~Park, H.S., Ryzhkov, A.V., Zrnic, D.S.: The hydrometeor classification algorithm for the polarimetric WSR-88DL Description and application to an MCS. Weather and Forecasting, 24, 730-748, 2009.~~

~~Reges, H.W., Doesken, N., Turner, J., Newman, N., Bergantino, A., Schwalbe, Z.: CoCoRaHS: The evolution and accomplishments of a volunteer rain gauge network. Bulletin of the American Meteorological Society, 97, 1831-1846, 2016.~~

~~Ruzanski, E., Chandrasekar, V.: Nowcasting rainfall fields derived from specific differential phase. Journal of Applied Meteorology and Climatology, 51, 1950-1959, 2012.~~

~~Ryzhkov, A.V., Giangrande, S., Schurr, T.: Rainfall measurements with the polarimetric WSR-88D radar. National Severe Storms Laboratory Rep. Norman: OK, 98, 2003.~~

~~Ryzhkov, A.V., Giangrande, S., Schurr, T.: Rainfall estimation with a polarimetric prototype of WSR-88D. Journal of Applied Meteorology, 44, 502–515, 2005.~~

~~Scarchilli, G., Gorgucci, E., Chandrasekar, V., Dobaie, A.: Self-consistency of polarization diversity measurement of rainfall. IEEE Transactions in Geoscience and Remote Sensing, 34, 22-26, 1996.~~

~~Shucksmith, P.E., Sutherland-Stacey, L., Austin, G.L.: The spatial and temporal sampling errors inherent in low resolution radar estimates of rainfall. Meteorological Applications, 18, 354-360, 2011.~~

~~Simpson, M.J., Hubbart, J.A., Fox, N.I.: Ground truthed performance of single and dual-polarized radar rain rates at large ranges. Hydrological Processes, 30, 3692-3703, 2016.~~

~~Simpson, M.J., Hirsch, A., Grempler, K., Lupo, A.R.: The importance of choosing precipitation datasets. Hydrological Processes, 1-13. doi: https://doi.org/10.1002/hyp.11381.~~

~~Seo, B.-C., Dolan, B., Krajewski, W., Rutledge, S.A., Petersen, W.: Comparison of single- and dual-polarization-based rainfall estimates using NEXRAD data for the NASA Iowa Flood Studies project. Journal of Hydrometeorology, 16, 1658-1675, 2015.~~

~~Smith, J.A., Seo, D.J., Baeck, M.L., Hudlow, M.D.: An intercomparison study of NEXRAD precipitation estimates. Water Resources Research, 32, 2035-2045, 1996.~~

~~Straka, J.M., Zrnic, D.S., Ryzhkov, A.V.: Bulk hydrometeor classification and quantification using polarimetric radar data: Synthesis of relations. Journal of Applied Meteorology, 39, 1341-1372, 2000.~~

~~Yang, L., Yang, Y., Liu, P., Wang, L.: Radar-derived quantitative precipitation estimation based on precipitation classification. Advances in Meteorology, 2016, 2016. doi: https://doi.org/10.1155/2016/2457489.~~

~~Zhang, G., Vivekanandan, J., Brandes, E.A.: A method for estimating rain rate and drop size distribution from polarimetric radar measurements. IEEE Transactions on Geoscience and Remote Sensing, 39, 830-841, 2001.~~

~~Zrnic, D.S., Ryzhkov, A.V.: Advantages of rain measurements using specific differential phase. Journal of Atmosphere and Oceanic Technology, 13, 454-464, 1996.~~

~~Zrnic, D.S., Ryzhkov, A.V.: Polarimetry for weather surveillance radars. Bulletin of American Meteorological Society, 80, 389-406, 1999.~~

~~Zrnic, D.S., Melknikov, V.M., Carter, J.K.: Calibrating differential reflectivity on the WSR-88D. Journal of Atmospheric and Oceanic Technology, 23, 944-951, 2005.~~

Formatted: Indent: First line: 0"

2.3 Radar data and radar-rainfall algorithms

Next Generation Radar (NEXRAD) level-II data were retrieved from the NCDC's HDSS. Files were processed using the Weather Decision Support System – Integrated Information (WDSS-II) program (Lakshmanan et al., 2007a) to assess reflectivity (Z) in addition to dual-polarized radar variables including differential reflectivity (ZDR) and specific differential phase shift (KDP). Three other variables were also generated based on a KDP-based smoothing field (Ryzhkov et al., 2003) for reflectivity, differential reflectivity, and specific differential phase: DSMZ, DZDR, and DKDP, respectively. These were implemented to determine whether the additional KDP-smoothing fields tend to over- or underestimate QPE's (Simpson et al., 2016). A rain rate echo classification variable (RREC) was also computed, which chooses whether an R(Z), R(KDP), R(Z,ZDR), or R(ZDR, KDP) algorithm is implemented in estimating rain rates based on the radar fields of Z, ZDR, and KDP (Kessinger et al., 2003) to determine whether a multi-parameter algorithm is superior to a single algorithm.

All seven variables (Z, ZDR, KDP, DSMZ, DZDR, DKDP, and RREC) were converted from their native polar grid to 256 x 256 1 km Cartesian grids, where the lowest radar elevation scans (0.5°) were used to mitigate uncalculated effects from evaporation and wind drift. An average of 5 minute scans were used for each of the variables, which were aggregated to hourly totals to be compared to the hourly tipping-bucket accumulations. In spite of previous reports suggesting 5 minute to hourly aggregates can have significant effects on QPE (e.g., Fabry et al. 1994), Shucksmith et al.'s (2011) criterion of present accumulation exceeding 26% for a pixel size of 1 km was not reached.

The latitude and longitude of each of the 15 gauges were matched with the radar pixel that corresponds to the Cartesian grid value of the seven radar variables which were then implemented in rain rate calculations. These rain-rate calculations were calculated using the equations presented by Ryzhkov et al. (2005) (Table 2), which were gathered from multiple studies using disdrometers to derive a relationship between reflectivity, differential reflectivity, and specific differential phase (Bringi and Chandrasekar, 2001; Brandes et al., 2002; Illingworth and Blackman, 2002; Ryzhkov et al., 2003). Standard R(Z) algorithms were also included to test whether the addition of dual-polarized technology improves QPE's.

With the use of both Z, ZDR, KDP, and DSMZ, DZDR, and DKDP fields produced by WDSS-II, the number of algorithms tested was 55. This includes the three standard single-polarized algorithms (stratiform, convective, and tropical) which were calculated using reflectivity R(Z), and then calculated as R(DSMZ), while algorithms 1-6 (R(KDP)) were also calculated as R(DKDP). Algorithms 7-11 (R(Z, ZDR)) were additionally calculated as R(Z, DZDR), R(DSMZ, ZDR), and R(DSMZ, DZDR), while the same four combinations of non- and KDP-smoothed fields were applied to the R(KDP, ZDR) algorithms (12-15). Quality controlling methods for the algorithms include mitigation of clutter, sun spikes, beam blockage, anomalous propagation, and removal of non-precipitation echoes (including biological and chaff returns) through w2qcnn the w2qcnndp algorithms (Lakshmanan et al., 2007b, 2010, 2014).

2.4 Statistical analyses

To test the performance of each algorithm, several statistical analyses were calculated. The average difference (Bias) was calculated as

Formatted: Indent: First line: 0.5"

$$Bias = \frac{\sum (R_i - G_i)}{N} \qquad (1)$$

Field Code Changed

where R_i is each hourly aggregated radar estimated rainfall amount calculated from one of the 55 algorithms, G_i is the hourly aggregated gauge (observed) measurement, and N is the total number of observations which, for this study, was 300 hours. A second statistical parameter, the normalized mean bias (NMB), was calculated as

$$NMB = \frac{1}{N} \frac{\sum (R_i - G_i)}{\sum G_i} \qquad (2)$$

Field Code Changed

The normalized mean bias is included in the analyses due to the fact that overestimations (i.e., radar estimates larger than gauge measurements) and underestimations (i.e., radar estimates smaller than gauge measurements) are treated proportionately. This is directly analogous to choosing the mean absolute error (MAE) opposed to the standard deviation as the MAE does not penalize smaller or larger errors, obscuring the overall results (Chai and Draxler, 2014). Bias measurements (Bias and NMB) were calculated to determine whether radar derived rain rates were over- or under-estimated in comparison to the gauges. However, to calculate the overall magnitude of error associated with the performance of the radars, the absolute values of (1) and (2) were performed to yield the mean absolute error (MAE), and normalized standard error (NSE), respectively.

Several other meteorological parameters were calculated, including probability of detection (PoD) which was calculated as

$$PoD = \frac{\sum | R_i \bullet G_i > 0 \,\&\, R_i > 0 |}{\sum | G_i |} \qquad (3)$$

Field Code Changed

where the bullet (•) indicates “if”, to determine how accurate the radars were at correctly detecting precipitation. The probability of detection values range between 0.0 (radar did not detect any precipitation

Field Code Changed

correctly) and 1.0 (radar detected the occurrence of all precipitation 100% correctly). The probability of false detection takes into account the amount of precipitation the radars incorrectly estimated when the gauges recorded zero values, and was calculated as

$$PoFD = \frac{\sum R_i \bullet (G_i = 0 \,\&\, R_i > 0)}{\sum G_i} \quad (4)$$

Quantitative measures including the missed precipitation amount (MPA) and the false precipitation amount (FPA) were defined such that

$$MPA = \sum R_i \bullet (G_i > 0 \,\&\, R_i = 0) \quad (5)$$

$$FPA = \sum R_i \bullet (G_i = 0 \,\&\, R_i > 0) \quad (6)$$

which analyzes the total amount of precipitation due to misses and false alarms. The total precipitation error was also recorded to assess the overall error from each radar.

3 Results and discussion

3.1 Overall algorithm performance

To test the overall performance of each radar, it was necessary to determine the overall best algorithm for each statistical measure. The best algorithm from each grouping of equations was determined to have the lowest normalized standard error (NSE), indicating the best performance relative to the gauge-recorded precipitation amount (Ryzhkov et al., 2005). This reduces the impact of bias inherent within the dataset between warm/cool season, stratiform/convective events, and allows for statistical measurements in spite of the (typical) non-Gaussian behavior of precipitation (Kleiber et al., 2012; Alaya et al., 2017).

From the results obtained, the three R(Z), three R(DSMZ), and RREC algorithms displayed a particular bias in favor of the R(Z)-Convective algorithm for all three radars with R(Z)-Stratiform displaying similar performance (Figure 2a). This could be due, at least in part, to the near-equal stratiform and convective precipitation regimes throughout 2014. Although errors generally increased as range increased for KEAX and KLSX, the results were nebulous for KSGF. The lowest NSE values were, typically, closest to each of the radars (between 0.4 and 0.8), with the notable exception of the closest gauge to KSGF. In general, the RREC performed worst at the largest of ranges, potentially due to the algorithm's ability to incorrectly assess the hydrometeors present (Cifelli et al., 2011; Yang et al. 2016). Additionally, the poor performance by the R(DSMZ)-Tropical equation is due to the lack of tropical precipitation within Central Missouri. Overall, the KDP-smoothed reflectivity fields (DSMZ) performed worse than their counter-parts, resulting in over-prediction of precipitation and, thus, larger errors (Simpson et al., 2016). Errors did not exceed 2.4 NSE units for any of these algorithms.

However, the performance of the KDP-smoothed KDP field (DKDP) performed better than the original specific differential phase shift field (Figure 2b). For nearly all gauges for each of the 3 radars, R(DKDP)4 performed the best, with NSE values ranging from 1.4 to 4.1. The range of NSE values were largest at KEAX, while the spread was relatively small for KLSX and KSGF. In spite of this, the overall spread of the performance of the 12 KDP algorithms varied greatly (average of 2 NSE units), exhibiting the sensitivity of KDP estimates on QPE (Ryzhkov et al., 2005; Cunha et al., 2013). In general, the NSSL-derived R(KDP) equations (i.e., equations 4-6) outperformed those from Bringi and Chandrasekar (2001, equation 1), Brandes et al. (2002, equation 2), and Illingworth and Blackman (2002, equation 3). Regardless, the magnitudes were all, approximately, more than 1 NSE unit than the performance of the R(Z) algorithms.

The algorithms with the lowest NSE values were equations 7-11. For example, the overall lowest NSE was at a distance of 130 km from KEAX (0.3), with no locations exceeding NSE values of 2.0 (Figure 2c). The large values at the closest location for KSGF (85 km, 1.3 – 1.9 NSE units), and the fifth closest gauge to KLSX (135 km, 1.3 – 1.8 NSE units), Cook Station, were similar to the R(Z) and R(DSMZ) results, indicating potential issues with reflectivity measurements. Additionally, these locations were the closest in performance to the R(KDP) and R(DKDP) NSE values. Observations from this gauge (Cook Station) indicated hail occurred during the evening of 01 August, for which KDP estimates would be more ideal than Z for QPE (Ryzhkov et al. 2005; Kumjian 2013a; Cunha et al. 2015). In spite of this, the overall spread in performance of the R(Z,ZDR) equations were less than the R(KDP) equations, demonstrating the robust performance of R(Z,ZDR) for QPE (Wang and Chandrasekar 2010; Seo et al., 2015).

The R(ZDR,KDP) algorithms performed the worst, overall (Figure 2d). In spite of the differential reflectivity being implemented, the overall NSE values increased in magnitude, exceeding 6 units for the second gauge analyzed by KEAX. Algorithms containing DKDP measurements performed better than simply KDP, demonstrating that even with the scaling behavior of ZDR, DKDP is superior to KDP estimates. This provides a potential solution to the noisy-ness that tends to be exhibited in the KDP field (Ruzanski and Chandrasekar 2012).

Due to the overall NSE values obtained, for the remainder of the analyses, equation 11 (i.e., R(Z,ZDR)5) and equation 13 (i.e., R(ZDR,KDP)2) will be utilized as the best and worst algorithms, respectively. Equations containing DZDR were not included in the following discussion due to the very large QPE errors for each radar.

3.2 KEAX

The overall bias showed that there was a positive bias, peaking near 5.5 mm hr^{-1} at the second gauge for KEAX, approximately 115 km from the radar for both the best and worst performing algorithms (Figure 3). This corresponds well with the spike in falsely detected precipitation recorded, which is canceled by the maximum in missed precipitation at the second distance of, approximately, 150 km. The overall worst algorithm, equation 13, an R(ZDR,KDP) relationship, revealed a decreasing trend in bias as the distance from the radar increased. For example, a bias of 4 mm hr^{-1} was observed at a distance of 75 km from the radar, whereas the bias reduced to 3 mm hr^{-1} at distances near 175 km. This could be due, at least in part, to the algorithm's utilization of KDP which performs poorly in frozen (especially light) precipitation (Zrnic and Ryzhkov, 1996; Kumjian 2013a), causing the overestimation. Conversely, the algorithm with the lowest bias was an R(Z,ZDR) algorithm (equation 11). There was a maximum in the bias calculations while utilizing equation 11 near 120 km, similar to equation 13, however, there was a more pronounced minimum in the data near 150 km. Furthermore, it appears the data oscillates around a bias value of 0 mm hr^{-1} when using equation 13. This could be due to ZDR's capability to respond to precipitation shape (Kumjian 2013a), which helps to scale the reflectivity portion of the rainfall estimation algorithm to a more accurate value (Seo et al., 2015). In general, the cool season displayed a larger magnitude of error in terms of bias for both algorithms.

The normalized mean bias (NMB) reveals the same trend in values for bias but with an overall decrease in magnitude. It is important to note, however, that the algorithms that tend to perform the worst (e.g., algorithms containing KDP) result in anomalous range responses which would be due, at least in part, to a stronger response to precipitation type. This indicates that observations above the melting layer are dominant for which QPE's tend not to be calculated (Cifelli et al., 2011; Seo et al., 2015) but are important for regions devoid of adequate radar coverage (Ryzhkov et al., 2003; Simpson et al., 2016).

The absolute bias and normalized standard error (NSE) shows the same maxima in the data at the second gauge (Brunswick) that was present in the bias data (6.2 mm hr^{-1} and 5.6, respectively) . However, a second maxima is located at the fifth gauge at, approximately, 150 km (Linneus) with values of 5.9 mm hr^{-1} and 4.0, respectively. Bright-band issues are detected due, at least in part, to the increased missed precipitation amount (240 mm) at this particular distance for the R(ZDR,KDP) equation (i.e., worst performing algorithm). There was also a pronounced minimum in the absolute bias and NSE results at the fourth gauge for equations 11 and 13, 4.0 mm hr^{-1} and 0.8 mm hr^{-1}, and 2.8 and 0.8, respectively, potentially indicating an idealized range of QPE for KEAX. Furthermore, the historical records at this particular gauge showed less issues (e.g., clogging) than any of the others analyzed by the KEAX radar. This highlights the importance of choosing ground-truth data, in particular tipping buckets which are prone to numerous errors (Ciach and Krajewski, 1999b).The largest contributions to the NSE and NMB were due to the warm season.

The probability of detection (PoD) results indicate a large difference in algorithm choice for correctly detecting precipitation. The low PoD at, approximately 150 km, indicates overshooting of the beam. This is further evidenced by the MPA results, as about 225 mm of precipitation was missed by the radar at 150 km, whereas only 100 mm of precipitation was missed by the radar at the second gauge at 120 km. Although equation 11, an R(Z,ZDR) algorithm was superior in terms of the bias, the same algorithm with a KDP-smoothed reflectivity value, R(DSMZ,ZDR) revealed the overall least amount of falsely missed precipitation (by 10 mm). However, the summation of the amount of precipitation falsely detected (PoFD) by KEAX showed a larger source of error than the MPA in terms of magnitude. For example, at the second (fifth) gauge, only 100 (225) mm of precipitation was missed by the radar, but over 700 (725) mm of precipitation was incorrectly estimated by the radar.

Correlation coefficient (CC) values for any of the 9 stations analyzed by KEAX ranges from 0.02 (Linneus, 151 km) to 0.93 for the cool season (St. Joseph, 115 km). The lowest R^2 were due to a combination of false alarms and misses. For example, the CC for the warm seasons at Sanborn (170 km) and Jefferson Farm (173 km) were 0.22 and 0.24, respectively, whereas when the instances of false alarms and misses were removed, increased to 0.48 and 0.52. Few locations (Brunswick, 114 km and Versailles, 129 km) saw little improvement in the CC values when only hits were analyzed (less than 0.1 increase), indicating the mean absolute error (in terms of hits) contributed the largest portion of error.

3.3 KLSX

Unlike the KEAX data, the gauges used for analyses for the KLSX radar span between 90 – 150 km. Furthermore, 5 out of the 8 gauges were located within 10 km of range from one-another, near 140 km from the radar, limiting the data available for analyses between 100 and 140 km (Figure 5).

The bias and NMB both show a relatively modest peak in values near the second gauge of 5 mm , which decreases to approximately 3.6 mm at the third gauge, 120 km from the radar. The worst performing algorithm, equation 13, was the same R(ZDR,KDP) relation as the worst KEAX bias and NMB data. Additionally, the overall trend of decreasing bias and NMB as distance from the radar increases was noted, presumably due to overshooting effects similar to the KEAX data. Furthermore, the overall non-biased results from the R(Z,ZDR) equation demonstrates its robust capabilities in QPE, in spite of its sensitivity to calibration (Zrnic et al., 2005; Bechini et al., 2008).

The double maxima in the absolute bias graph are present as with the KEAX data, but are not as pronounced. For example, the absolute bias at 95 km and 140 km from KLSX were 5.9 mm and 1.1 mm , and 4.9 mm and 1.4 mm for equations 13 and 11, respectively. Additionally, the overall minima in the absolute bias for both KEAX and KLSX are at, approximately, 125 km from the radar (3.9 mm hr^{-1} and 1.0 mm hr^{-1}, respectively, for equations 13 and 11). The relative distance from the radars are the same, where the two maxima for KEAX were at 115 and 150 km, while the maxima were at, approximately, 100 and 140 km for KLSX. The overall best and worst performing algorithms at KLSX for the absolute bias and NSE were equations 11 and 13, the R(Z,ZDR) and R(ZDR,KDP) algorithms, respectively.

The magnitude of error in terms of absolute bias, normalized mean bias, and normalized standard error, all showed a decreasing pattern as distance from KLSX increased. This was due, primarily, from a maximum in the false precipitation amount at 95 km from the radar. Historical notes at this location indicate frequent clogging of the rain gauge, either due to bugs or leaves. From a particular series of events spanning from 01 to 04 April and 01 to 03 August, 2014, over 130 mm of precipitation occurred during each period which was not captured by the gauge, resulting in a large amount of overall error. These results indicate the important of dual gauges in the same vicinity (Krajewski et al. 1998; Ciach and

Krajewski 1999). Interestingly, the cool season displayed a larger NSE (5 % for R(ZDR,KDP)) potentially due to the very low probability of detection (0.2) at this range of 118 km.

One of the main differences between the KLSX and KEAX data was the decreased probability of detection at 120 km for KLSX, while there was an increased probability of detection for KEAX. In general, the PoD values were worse for KLSX when compared to KEAX. For example, equation 11 had no PoD values below 0.90, whereas no PoD values exceeded 0.84 for KLSX. There was also a slight trend of increasing PoD values as distance from the St. Louis radar increased and, at one point near 140 km, the best algorithm, R(DSMZ) convective and the worst algorithm, KDP1, were not significantly different ($p < 0.10$). Additionally, the maxima in the PoD while utilizing KDP1 corresponds to a minima in the R(DSMZ) detection percentage, which is well correlated by the similarly valued MPA results.

The missed precipitation amount (MPA) displayed the cool season contributed the most, whereas the warm season contributed the most amount of false precipitation amount. The R(Z,ZDR) equation only registered, on average, 25 mm of MPA and 160 mm of FPA, whereas the R(ZDR,KDP) equation was very dependent upon range. For example, the FPA from R(ZDR,KDP) decreased as range increased from the radar from a maximum of, approximately, 850 mm to 620 mm. However, the fifth-furthest gauge (137 km from KLSX) displayed a sharp increase in the MPA for both cool seasons (above 100 mm).

3.4 KSGF

In spite that the KLSX and KEAX data strongly suggests false precipitation errors near 100 km in addition to bright-banding near 150 km from the radars, the KSGF results reveal an overall smooth decrease (increase) of error with range (Figure 7) for R(ZDR,KDP) and R(Z,ZDR), accordingly. One of the main reasons for this could be due to the fact that only 5 gauges were analyzed from KSGF (the fewest of the 3 radars analyzed), smoothing the overall trend lines.

The bias remained relatively constant near -0.3 mm for R(Z,ZDR), whereas the bias exhibited a sharp decrease from 4 mm to 2.7 mm over a distance of, approximately, 100 km. In general, the cool season displayed the lower of bias magnitudes when compared to the warm season, similar to the KEAX results. This may be due, at least in part, to the low PoFD values for the warm season close to the KSGF radar.

Similar to the bias, the absolute bias for R(Z,ZDR) was constant at all ranges (near 1 mm) whereas the R(ZDR,KDP) equation decreased from 5.2 mm to 3.8 mm. This is potentially due to the low cool season PoD values (below 0.6), while the warm season R(ZDR,KDP) values (near 0.8) remained constant. A larger contribution from more correctly detected precipitation in addition to the decreasing trends in the NMB and NSE would result in a lower absolute bias.

The closest location (90 km) typically displayed the largest errors for the R(ZDR,KDP) equation, and then decreased in error magnitude as range increased. In spite of this, the PoFD results indicate both algorithms increased in PoFD values as range increased, with the warm season typically dominating, particularly due to the large convective clouds dominate in the warm season. False detection values as low as 0.01 for the cool season while utilizing R(Z,ZDR) were observed at distances near 100 km and 140 km from the radar.

Normalized standard error values increased from 0.7 % at a distance of 105 km to 1.8 % at a distance of 185 km for R(Z,ZDR). Large NSE values for the warm season (7.5 %) were calculated for R(ZDR,KDP) which decreased to 3.8 % at 185 km from the radar. Furthermore, this was the only instance when the warm season was less than the cool season in terms of NSE. Otherwise, the overall NSE decreased from 5 % to 3.9 % for R(ZDR,KDP). The NMB followed a similar trend for the KDP-containing algorithm, with a noticeable exception at the second gauge (105 km from KSGF), where the overall NSE was closer to the warm than cool season. This is due to the low PoFD values at this location, in addition to a smaller difference between the two algorithm’s FPA measurements.

The MPA results, unlike for KEAX and KLSX, displayed a larger range of performance between seasons. However, the warm season still exhibited the overall best performance in terms of MPA, yet contributed the most to the FPA for both R(Z,ZDR) and R(ZDR,KDP). In spite of the MPA typically increasing as range increased, the FPA was more nebulous. For example, the second gauge (105 km from KSGF) had the overall lowest NSE (0.8 %), MPA (15 mm), and FPA (95 mm) for R(Z,ZDR). The third-furthest location (142 km) resulted in slightly larger errors, overall, while the fourth-furthest location had errors similar to the second gauge (105 km). Then, at the furthest tipping bucket location (185 km), NSE values increased, whereas FPA and MPA decreased. Therefore, the furthest location's errors are due, primarily, from discrepancies between precipitation magnitude between the gauge and radar.

Excluding Versailles (142 km from KSGF), the cool season exhibited larger R^2 values in comparison to the cool season (Figure 8). Furthermore, CC values exceeded 0.9 when false alarms and misses were excluded from Mt. Grove (101 km) and was 0.84 when included. Otherwise, the other four stations analyzed by the Springfield radar displayed many counts of false alarms and misses, leading to low R^2 values.

Due to the relatively large ranges from the Springfield (KSGF) radar, most of the correlation coefficient values were low in comparison to either KLSX or KEAX. For the warm (cool) season without false alarms and misses, R^2 values ranged from 0.44 (0.38) and 0.34 (0.36) for KLSX and KSGF, respectively, at Cook Station (119 and 185 km). Similarly, the CC values ranged from 0.61 (0.71) to 0.42 (0.56) at Green Ridge (76 and 154 km) for KEAX and KSGF, accordingly.

4 Conclusions

Dual-polarization technology was implemented to the National Weather Service Next Generation Radar network in the Spring of 2012 to, primarily, improve quantitative precipitation estimation and hydrometeor classification. The current study observed over 300 hours of precipitation data with three separate radars in Missouri using 55 algorithms including the three conventional R(Z) radar rain-rate estimation algorithms (stratiform, convective, and tropical) along with a myriad of R(KDP), R(Z,ZDR), and R(ZDR,KDP) algorithms which can be found in Ryzhkov et al. (2005). Additionally, a KDP-smoothing field of reflectivity, differential reflectivity, and the specific differential phase shift (DSMZ, DZDR, and DKDP, respectively) were measured and used for analyses. Unlike previous studies, the current work emphasizes the amount of precipitation correctly and incorrectly estimated by the radar in comparison to the terrestrial based precipitation gauges through measurements of the missed and false precipitation amount.

For all three radars, Kansas City, St. Louis, and Springfield, MO (KEAX, KLSX, and KSGF, respectively), the majority of precipitation error (over 60%) was contributed by the amount of precipitation falsely detection by the radar (up to 725 mm), while 20% was due to the radar missing the precipitation (up to 225 mm) for KEAX. Similar magnitudes of error were reported for KLSX and KSGF, with an overall error in precipitation for each radar ranging between 250 mm for the best performing of the 55 algorithms, equation 11 (an R(Z,ZDR) algorithm), and up to 2000 mm for the worst performing algorithms, R(ZDR,KDP) equation 13. The R(Z,ZDR) equation (an NSSL algorithm) was determined to be the most robust due to it registering the lowest NSE. These values of false precipitation amount and missed precipitation amount generally increased as range from the radar increased.

Most algorithms showed a degradation in the normalized standard error with range. In particular, the KDP-smoothed equations displayed larger biases and NSE values than their non-KDP counterparts, with the exception of R(KDP) algorithms themselves. Some larger errors were recorded at gauge locations close to the radar, potentially due to bright-banding effects which were determined to be due to the large false precipitation amount analyzed at these locations.

The data was divided into summer (May – October) and winter (November – April; 59 and 41% of the entire data, respectively). Despite the winter data contributing less than the summertime data, it accounted for 20% of the overall MPA, and 40% to the overall PoFD. The R^2 values were less during the winter in comparison to the warm season primarily due to the smaller magnitude of precipitation that occurred. Furthermore, CC values increased by as much as 0.4 when instances of hits and misses were removed from the analyses, resulting in the warm season to outperform the cool season CC values at particularly short ranges from the radar.

These results aid in our understanding in the possibilities for hydrometeorological studies. Nearly 50% of the 300 hours where precipitation occurred analyzed for the study consisted of either falsely estimated precipitation by the radar, or missed by the radar. Furthermore, these errors accumulate between 500 to 2,000 mm of precipitation depending on the algorithms chosen. Although the overall performance increased when false alarms and misses were removed, correlation coefficient values still, typically, remained below 0.50 at ranges beyond 130 km.

Furthermore, results demonstrate the issues with analyzing QPE from a single gauge, explaining why the Community Collaborative Rain, Hail, and Snow Network (Kelsch 1998; Cifelli et al., 2005; Reges et al., 2016) or other densely-gauged networks (e.g., the Hydrometeorological Automated Data System, HADS, Meteorological Assimilation Data Ingest System, MADIS) tends to be more utilized since results have shown that measurements or quality controlled-techniques made by these organizations, especially CoCoRaHS, are significantly more accurate than rain gauges (Simpson et al., 2017), especially for convective events (Moon et al. 2009).

Author Contribution. N. Fox designed the experiment and provided feedback while M. Simpson carried out the calculations and wrote the manuscript.

Acknowledgements. This material is based upon work supported by the National Science Foundation under Award Number IIA-1355406. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Alaya, M.A., Ourda, T.B.M.J., Chebana, F.: Non-Gaussian spatiotemporal simulation of multisite precipitation: Downscaling framework. Climate Dynamics, 2017. doi: https://doi.org/10.1007/s00382-017-3578-0.

Anagnostou, M.N., Anagnostou, E.N., Vulpiani, G., Montopoli, M., Marzano, F.S., Vivekanandan, J.: Evaluation of X-band polarimetric-radar estimates of drop-size distributions from coincident S-band polarimetric estimated and measured raindrop spectra. IEEE Transactions on Geoscience and Remote Sensing, 46, 3067-3075, 2008.

Bechini, R., Baldini, L., Cremonini, R., Gorgucci, E.: Differential reflectivity calibration for operational radars. Journal of Atmospheric and Oceanic Technology, 25, 1542-1555, 2009.

Berne, A. and Uijlenhoet, R.: A stochastic model of range profiles of raindrop size distributions: application to radar attenuation correction, Geophysical Research Letters, 32, 2005, doi: https://doi.org/10.1029/2004GL021899.

Berne, A. and Krajewski, W.F.: Radar for hydrology: Unfulfilled promise or unrecognized potential? Advances in Water Resources, 51, 357-366, 2013.

Bringi, V.N. and Chandrasekar, V.: Polarimetric Doppler weather radar, principles and applications. Cambridge University Press: Cambridge, UK, 636, 2001.

Brandes, E.A., Zhang, G., Vivekanandan, J.: Experiments in rainfall estimation with a polarimetric radar in a subtropical environment, Journal of Applied Meteorology, 41, 674–685, 2002.

Brandes, E.A., Zhang, G., Vivekanandan, J.: Drop size distribution retrieval with polarimetric radar: model and application, Journal of Applied Meteorology, 43, 461-475, 2004.

Chai, T., Draxler, R.R.: Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature, Geoscientific Model Development, 7, 1247-1250, 2014.

Ciach, G.J., Krajewski, W.F.: On the estimation of radar rainfall error variance. Advances in Water Resources, 22, 585-595, 1999a.

Ciach, G.J. and Krajewski, W.F.: Radar-raingage comparisons under observational uncertainties. Journal of Applied Meteorology, 38, 1519-1525, 1999b.

Ciach, G.J.: Local random errors in tipping-bucket rain gauge measurements. Journal of Atmospheric and Oceanic Technology, 20, 752-759, 2002.

Cifelli, R., Doesken, N., Kennedy, P., Carey, L.S., Rutledge, S.A., Gimmestad, C., Depue, T.: The community collaborative rain, hail, and snow network: Informal education for scientists and citizens. Bulletin of the American Meteorological Society, 86, 1069-1077, 2005.

Cunha, L.K., Smith, J.A., Baeck, M.L., Krajewski, W.F.: An early performance of the NEXRAD dual-polarization radar rainfall estimates for urban flood applications. Weather and Forecasting, 28, 1478-1497, 2013.

Cunha, L.K., Smith, J.A., Krajewski, W.F., Baeck, M.L., Seo, B.: NEXRAD NWS polarimetric precipitation product evaluation for IFloods. Journal of Hydrometeorology, 16, 1676-1699, 2015.

Delrieu, G., Andrieu, H., Creutin, J.D.: Quantification of path-integrated attenuation for X- and C-band weather radar systems operating in Mediterranean heavy rainfall. Journal of Applied Meteorology, 39, 840-850, 2000.

Fabry, F., Bellon, A., Duncan, M.R., Austin, G.L.: High resolution rainfall measurements by radar for very small basins: the sampling problem reexamined. Journal of Hydrology, 161, 415-428, 1994.

Fairman, J.G., Schultz, D.M., Kirschbaum, D.J., Gray, S.L., Barrett, A.I.: A radar-based rainfall climatology of Great Britain and Ireland. Weather, 70, 153-158, 2012. doi: https://doi.org/10.1002/wea.2486.

Gamache, J.F. and Houze, R.A.: Mesoscale air motions associated with a tropical squall line. Monthly Weather Review, 110, 118–135, 1982.

Giangrande, S.E. and Ryzhkov, A.V.: Estimation of rainfall based on the results of polarimetric echo classification. Journal of Applied Meteorology, 47, 2445-2460, 2008.

Gorgucci, E., Scarchilli, G., Chandrasekar, V.: Calibration of radars using polarimetric techniques. IEEE Transactions in Geoscience and Remote Sensing, 30, 853-858, 1992.Gorgucci, E., Scarschilli, G., Chandrasekar, V., Bringi, V.N.: Measurement of mean raindrop shape from polarimetric radar observations. Journal of the Atmospheric Sciences, 57, 3406-3413, 2000.

Gorgucci, E., Baldini, L., Chandrasekar, V.: What is the shape of a raindrop? An answer from radar measurements. Journal of the Atmospheric Sciences, 63, 3033-3044, 2006.

Goudenhoofdt, E., Delobbe, L.: Long-term evaluation of radar QPE using VPR correction and radar-gauge merging. International Association of Hydrological Sciences Publications, 351, 249-254, 2012.

Goudenhoofdt, E., Delobbe, L.: Generation and verification of rainfall estimates from 10-yr volumetric weather radar measurements. Journal of Hydrometeorology, 133, 1191-1204, 2016.

Gourley, J.J., Giangrande, S.E., Hong, Y., Flamig, Z., Schuur, T., Vrugt, J.: Impacts of polarimetric radar observations on hydrologic simulation. Journal of Hydrometeorology, 11, 781-796, 2010.

Habib, E., Krajewski, W.F., Nespor, V., Kruger, A.: Numerical simulation studies of rain gauge data correction due to wind effect. Journal of Geophysical Research, 104, 723–734, 1999.

Habib, E., Krajewski, W.F., Kruger, A.: Sampling errors of tipping-bucket rain gauge measurements. Journal of Hydrological Engineering, 6, 159–166, 2001.

Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D., New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. Journal of Geophysical Research, 113, 2008. doi: https://doi.org/10.1029/2008JD010201

Holleman, I., Huuskonen, A., Gill, R., Tabary, P.: Operational monitoring of radar differential reflectivity using the sun. Journal of Atmospheric and Oceanic Technology, 27, 881-887, 2010.

Hubbert, J.C.: Differential reflectivity calibration and antenna temperature. Journal of Atmospheric and Oceanic Technology, 34, 1885-1906, 2017.

Illingworth, A., Blackman, T.A.: The need to represent raindrop size spectra as normalized gamma distributions for the interpretation of polarization radar observations. Journal of Applied Meteorology, 41, 286-297, 2002.

Kelsch, M.: The Fort Collins flash flood: Exceptional rainfall and urban runoff. Preprints, 19th Conference on severe local storms, Minneapolis, MN, American Meteorological Society, 404-407, 1998.

Kitchen, M. and Blackall, M.: Representativeness errors in comparisons between radar and gauge measurements of rainfall. Journal of Hydrology, 134, 13–33, 1992.

Kleiber, W., Katz, R.W., Rajagopalan, B.: Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. Water Resources Research, 48, 2012. doi: https://doi.org/10.1029/2011WR011105.Kessinger, C., Ellis, S., Van Andel, J.: The radar echo classifier: a fuzzy logic algorithm for theWSR-88D. 19th Conf. on Inter. Inf. Proc. Sys. (IIPS) for Meteor., Ocean., and Hydr., Amer. Meteor. Soc., Long Beach, CA, 2003.

Kitchen, M. and Jackson, P.M.: Weather radar performance at long range – simulated and observed. Journal of Applied Meteorology, 32, 975-985, 1993.

Krajewski, W.F., Kruger, A., Nespor, V.: Experimental and numerical studies of small-scale rainfall measurements and variability. Water Science and Technology, 37, 131-138.

Kumjian, M.R.: Principles and applications of dual-poarization weather radar. Part 1: Description of the polarimetric radar variables. Journal of Operational Meteorology, 1, 226-242, 2013a.

Kumjian, M.R.: Principles and applications of dual-poarization weather radar. Part 2: Warm and cold season applications. Journal of Operational Meteorology, 1, 243-264, 2013b.

Kumjian, M.R.: Principles and applications of dual-poarization weather radar. Part 3: Artifacts. Journal of Operational Meteorology, 1, 265-274, 2013c.

Lakshmanan, V., Smith, T., Stumpf, G., Hondl, K.: The warning decision support system—integrated information. Weather and Forecasting, 22, 596–612, 2007a.

Lakshmanan, V., Fritz, A., Smith, T., Hondl, K., Stumpf, G.: An automated technique to quality control radar reflectivity data. Journal of Applied Meteorology and Climatology, 46, 288-305, 2007b.

Lakshmanan, V., Zhang, J., Howard, K.: A technique to censor biological echoes in radar reflectivity data. Journal of Applied Meteorology and Climatology, 49, 453-462, 2010.

Lakshmanan, V., Karstens, C., Krause, J., Tang, L,: Quality control of weather radar data using polarimetric variables. Journal of Atmospheric and Oceanic Technology, 31, 1234-1249, 2014.

Moon, J.T., Guinan, P.E., Snider, D.J., Lupo, A.R.: CoCoRaHS in Missouri: Four years later, the importance of observations. Transactions of the Missouri Academy of Science, 43, 7-18, 2009.

Park, H.S., Ryzhkov, A.V., Zrnic, D.S.: The hydrometeor classification algorithm for the polarimetric WSR-88DL Description and application to an MCS. Weather and Forecasting, 24, 730-748, 2009.

Reges, H.W., Doesken, N., Turner, J., Newman, N., Bergantino, A., Schwalbe, Z.: CoCoRaHS: The evolution and accomplishments of a volunteer rain gauge network. Bulletin of the American Meteorological Society, 97, 1831-1846, 2016.

Ruzanski, E., Chandrasekar, V.: Nowcasting rainfall fields derived from specific differential phase. Journal of Applied Meteorology and Climatology, 51, 1950-1959, 2012.

Ryzhkov, A.V., Giangrande, S., Schurr, T.: Rainfall measurements with the polarimetric WSR-88D radar. National Severe Storms Laboratory Rep. Norman: OK, 98, 2003.

Ryzhkov, A.V., Giangrande, S., Schurr, T.: Rainfall estimation with a polarimetric prototype of WSR-88D. Journal of Applied Meteorology, 44, 502–515, 2005.

Scarchilli, G., Gorgucci, E., Chandrasekar, V., Dobaie, A.: Self-consistency of polarization diversity measurement of rainfall. IEEE Transactions in Geoscience and Remote Sensing, 34, 22-26, 1996.

Shucksmith, P.E., Sutherland-Stacey, L., Austin, G.L.: The spatial and temporal sampling errors inherent in low resolution radar estimates of rainfall. Meteorological Applications, 18, 354-360, 2011.

Simpson, M.J., Hubbart, J.A., Fox, N.I.: Ground truthed performance of single and dual-polarized radar rain rates at large ranges. Hydrological Processes, 30, 3692-3703, 2016.

Simpson, M.J., Hirsch, A., Grempler, K., Lupo, A.R.: The importance of choosing precipitation datasets. Hydrological Processes, 1-13. doi: https://doi.org/10.1002/hyp.11381.

Seo, B.-C., Dolan, B., Krajewski, W., Rutledge, S.A., Petersen, W.: Comparison of single- and dual-polarization-based rainfall estimates using NEXRAD data for the NASA Iowa Flood Studies project. Journal of Hydrometeorology, 16, 1658-1675, 2015.

Smith, J.A., Seo, D.J., Baeck, M.L., Hudlow, M.D.: An intercomparison study of NEXRAD precipitation estimates. Water Resources Research, 32, 2035-2045, 1996.

Straka, J.M., Zrnic, D.S., Ryzhkov, A.V.: Bulk hydrometeor classification and quantification using polarimetric radar data: Synthesis of relations. Journal of Applied Meteorology, 39, 1341-1372, 2000.

Yang, L., Yang, Y., Liu, P., Wang, L.: Radar-derived quantitative precipitation estimation based on precipitation classification. Advances in Meteorology, 2016, 2016. doi: https://doi.org/10.1155/2016/2457489.

Zhang, G., Vivekanandan, J., Brandes, E.A.: A method for estimating rain rate and drop size distribution from polarimetric radar measurements. IEEE Transactions on Geoscience and Remote Sensing, 39, 830-841, 2001.

Zhang, J., Youcun, Q.: A real-time algorithm for the correction of brightband effects in radar-derived QPE. Journal of Hydrometeorology, 11, 1157-1171.

Zhang, J., Langston, C., Howard, K.: Brightband identification based on vertical profiles of reflectivity from the WSR-88D. Journal of Atmospheric and Oceanic Technology, 25, 1859-1872.

Zrnic, D.S., Ryzhkov, A.V.: Advantages of rain measurements using specific differential phase. Journal of Atmosphere and Oceanic Technology, 13, 454-464, 1996.

Zrnic, D.S., Ryzhkov, A.V.: Polarimetry for weather surveillance radars. Bulletin of American Meteorological Society, 80, 389-406, 1999.

Zrnic, D.S., Melknikov, V.M., Carter, J.K.: Calibrating differential reflectivity on the WSR-88D. Journal of Atmospheric and Oceanic Technology, 23, 944-951, 2005.

Table 1. Terrestrial-based precipitation gauge locations used for the study in addition to the National Weather Service Radars Springfield, MO (KSGF), Kansas City, MO (KEAX), and St. Louis, MO (KLSX) used in conjunction with each gauge.

Gauge Location	Latitude (°N)	Longitude (°W)	Radar(s) Used
Bradford	38.897236	-92.218070	KLSX, KEAX
Brunswick	39.412667	-93.196500	KEAX
Capen Park	38.929237	-92.321297	KLSX, KEAX
Cook Station	37.797945	-91.429645	KLSX, KSGF
Green Ridge	38.621147	-93.416652	KEAX, KSGF
Jefferson Farm	38.906992	-92.269976	KLSX, KEAX
Lamar	37.493366	-94.318185	KSGF
Linneus	39.856919	-93.149726	KEAX

Monroe City	39.635314	-91.725370	KLSX
Mountain Grove	37.153865	-92.268831	KSGF
Sanborn Field	38.942301	-92.320395	KLSX, KEAX
St. Joseph	39.757821	-94.794567	KEAX
Vandalia	39.302300	-91.513000	KLSX
Versailles	38.434700	-92.853733	KEAX, KSGF
Williamsburg	38.907350	-91.734210	KLSX

Table 2. List of single- and dual-polarimetric algorithms used for radar rainfall estimates.

$R(Z) = aZ^b$			
Precipitation type	a	b	c
Stratiform	200	1.6	-
Convective	300	1.4	-
Tropical	250	1.2	-
$R(KDP) = a \mid KDP \mid^b \; sign(KDP)$			
Algorithm number			
1	50.7	0.85	-
2	54.3	0.81	-
3	51.6	0.71	-

4	44.0	0.82	-
5	50.3	0.81	-
6	47.3	0.79	-
$R(Z, ZDR) = aZ^{b}ZDR^{c}$			
Algorithm number			
7	6.70×10^{-3}	0.927	-3.43
8	7.46×10^{-3}	0.945	-4.76
9	1.42×10^{-2}	0.770	-1.67
10	1.59×10^{-2}	0.737	-1.03
11	1.44×10^{-2}	0.761	-1.51
$R(ZDR, KDP) = a\,\lvert KDP\rvert^{b}\ ZDR^{c} sign(KDP)$			
Algorithm number			
12	90.8	0.930	-1.69
13	136	0.968	-2.86
14	52.9	0.852	-0.53
15	63.3	0.851	-0.72

Figures

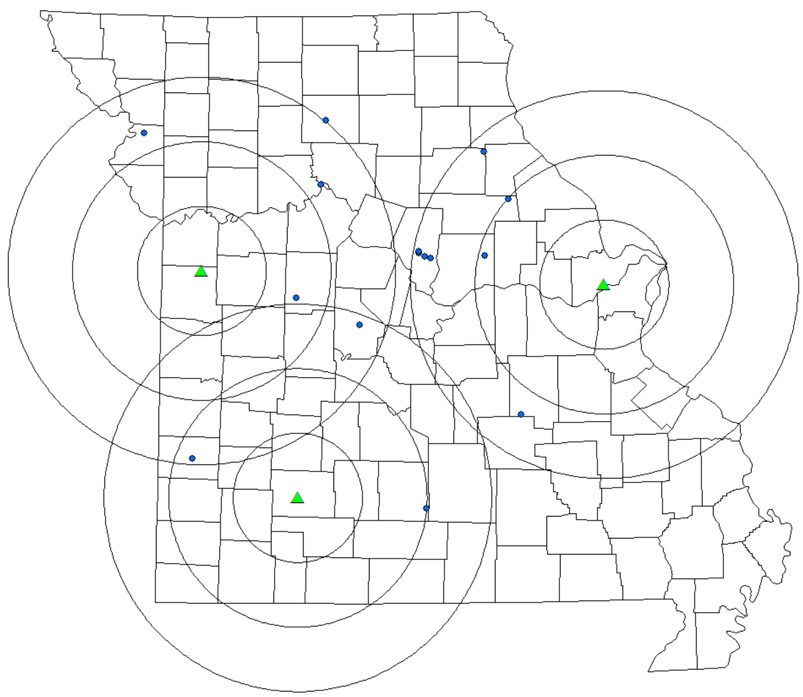

Figure 1. Study location (Missouri) with St. Louis (KLSX), Kansas City (KEAX), and Springfield (KSGF), MO radars (triangles) overlaid with 50-, 100-, and 150-km range rings in addition to the 15 terrestrial-based precipitation gauges utilizeed as ground-truthed data.

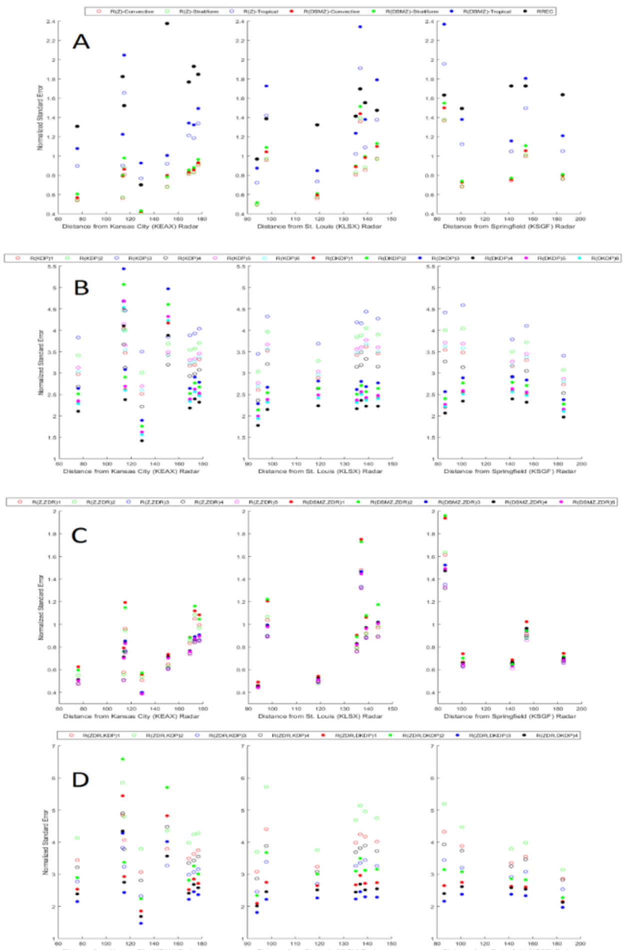

Figure 2. Normalized standard error values for the overall performance of the a) 3 R(Z), 3 R(DSMZ), and RREC algorithms, b) 6 R(KDP) and 6 R(DKDP) algorithms (equations 1-6 from Table 2), c) 5 R(Z,ZDR) and 5 R(DSMZ,ZDR) algorithms (equations 7-11 from Table 2), and d) 4 R(ZDR,KDP) and 4 R(ZDR,DKDP) algorithms (equations 12-15 from Table 2) for the three radars utilized for the current study.

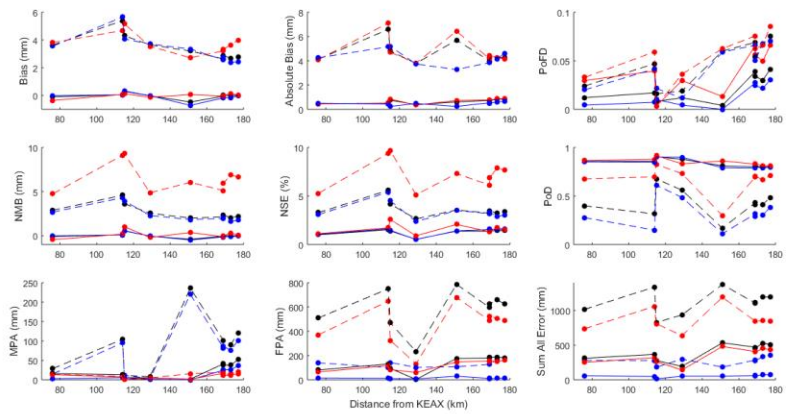

Figure 3. Values of analyses from the Kansas City (KEAX) radar. Dashed lines and points represent the analyses of the worst-performing algorithm (R(ZDR,KDP)) while the solid lines and points represent the analyses of the best-performing algorithm (R(Z,ZDR)). Red, blue, and black colors represent analyses conducted during the warm and cool seasons, and overall, respectively.

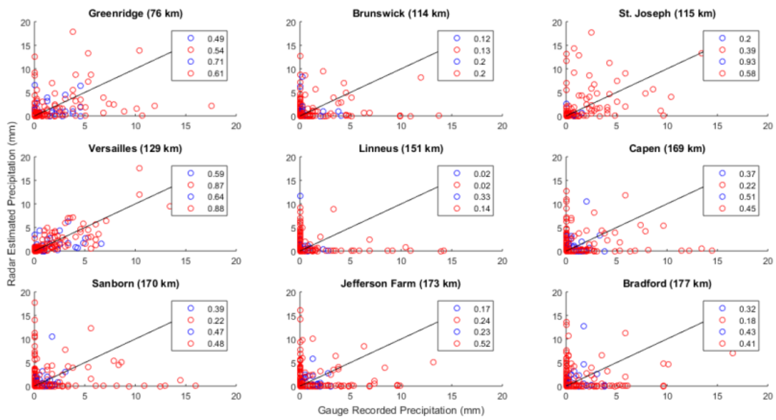

Figure 4. Correlation coefficient values for the 9 locations analyzed by the Kansas City (KEAX) radar with the R(Z,ZDR) NSSL equation. Blue and red scatter points represent the cool and warm season data, respectively. The top two numbers on each plot indicate the overall R^2 value, whereas the bottom two numbers represent the R^2 when false alarms and misses are removed.

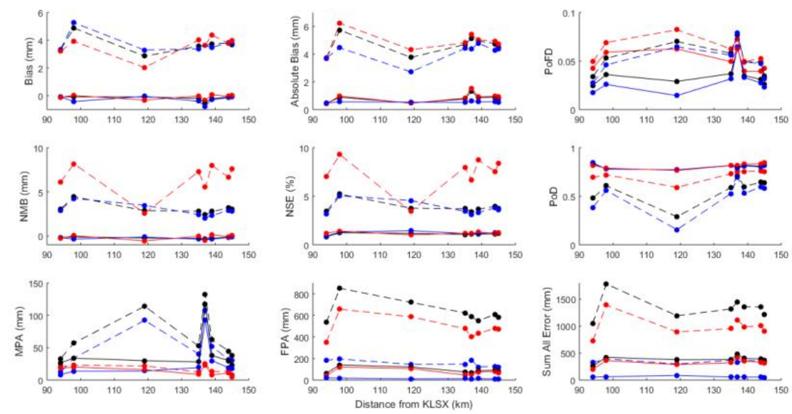

Figure 5. Values of analyses from the St. Louis (KLSX) radar. Dashed lines and points represent the analyses of the worst-performing algorithm (R(ZDR,KDP)) while the solid lines and points represent the analyses of the best-performing algorithm (R(Z,ZDR)). Red, blue, and black colors represent analyses conducted during the warm and cool seasons, and overall, respectively.

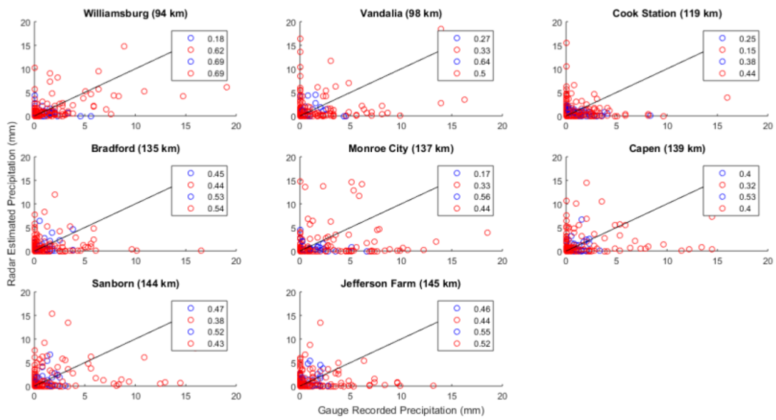

Figure 6. Correlation coefficient values for the 8 locations analyzed by the St. Louis (KLSX) radar with the R(Z,ZDR) NSSL equation. Blue and red scatter points represent the cool and warm season data, respectively. The top two numbers on each plot indicate the overall R^2 value, whereas the bottom two numbers represent the R^2 when false alarms and misses are removed.

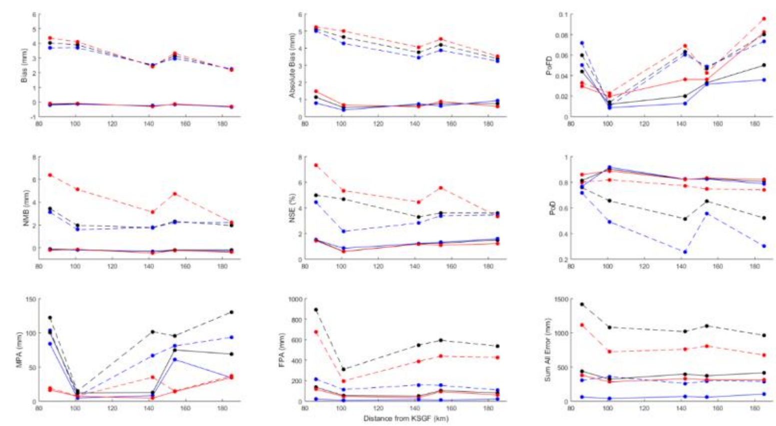

Figure 7. Values of analyses from the Springfield (KSGF) radar. Dashed lines and points represent the analyses of the worst-performing algorithm (R(ZDR,KDP)) while the solid lines and points represent the analyses of the best-performing algorithm (R(Z,ZDR)). Red, blue, and black colors represent analyses conducted during the warm and cool seasons, and overall, respectively.

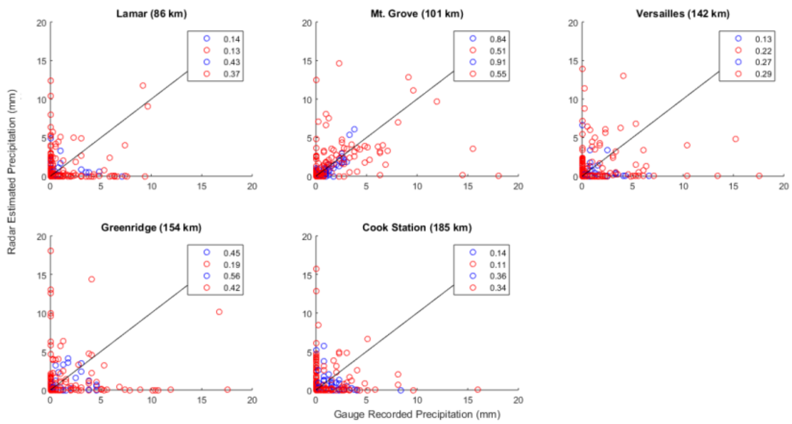

Figure 8. Correlation coefficient values for the 5 locations analyzed by the Springfield (KSGF) radar with the R(Z,ZDR) NSSL equation. Blue and red scatter points represent the cool and warm season data, respectively. The top two numbers on each plot indicate the overall R^2 value, whereas the bottom two numbers represent the R^2 when false alarms and misses are removed.