First, we wish to express our thanks to the anonymous reviewer and to Hidde Leijnse for their efforts and constructive comments. As both reviewers suggested, we have rewritten the last part of the paper, changing its title from "summary" to "conclusions" while discussing the major lessons derived from our study. Detailed answers to the reviewers' comments are given below.

Reviewer #1

General Comments

I looked forward to reading the paper and expected reviewing it to be a breeze but it turned into a chore. I found many little points that need to be corrected before I can recommend the paper for publication. Overall the paper is interesting and adds to our knowledge of rainfall variability. It combines an observational study with data analyses. The authors are clearly familiar with relevant findings in the literature but add little new to the data analysis (however, this is not a condition for this paper to be publishable). I hope that the authors will continue running the network, perhaps even expanding it, and that in the future when the data sample grows larger, they will share many more insightful results with the radar hydrology community.

Specific Comments

Let's start with the title. Was there just one lesson learned? I don't think so. Recommend changing to "lessons learned".
The title was changed as the reviewer suggested.

Abstract

Line 9. "Pixel" refers to an image or a field or a map. The authors talk about sub-pixel before defining the size of pixel.
Here, "pixel" refers to the remote-sensing data dimension. Two sentences were revised (this, and line 6) and information regarding the time scales used was added (see our next answer).
"This network was established for a detailed exploration of the uncertainties and errors regarding rainfall variability within a common pixel size of data obtained from remote sensing systems **for time scales of 1-min to daily** ".

Line 14. Up to this point in the abstract the authors said nothing about the time scales they investigated.
The reviewer is right, it was corrected in line 8 (see previous correction).

Line 13. When you say "zero-distance correlation. . ." the readers don't know what you are talking about. You need to mention first that your network has 13 stations, each with two side-by-side gauges.
Right! It was added to the abstract as follows (line 6): "… a new super-dense network of rain gauges containing **13 stations, each with two side-by-side gauges**, was installed **within a 4 km^2 study area** near Kibbutz Galed in northern Israel".

Line 14-16. This sentence is technically incorrect and incomprehensible to those who are not already familiar with the VRF concept.
This sentence was rewritten: "The variance reduction factor (VRF), representing the uncertainty from averaging a number of rain stations per pixel, ranged from 1.6% for the 1-min time scale to 0.07% for the daily scale".

Line 16-17. This statement is meaningless without some criterion for "representation". . .

Right, it is explained in page 12, line 10 that we looked for value of VRF<5%. It was also added to the abstract (see the revised abstract text below).

Line 17-19. “The radar-rain gauge error. . .” What is that? I suspect, but only because I am familiar with the problem, that you wanted to say something to the effect of “Radar rain vs. gauge-rain difference” (not error).
We changed the terminology used here (and further on) from "error" to "difference".

Line 20. In this business the truth is unknown and can only be estimated, better or worse. You should not use terms like “ground truth” without a clear context.
The true rainfall is indeed unknown- therefore we used the areal-averaged gauge-derived rainfall (as explained in section 5.2). As the reader may be confused we changed this sentence to: "The ratio of radar rainfall to gauge-areal-averaged rainfall, expressed by the error distribution scatter parameter, decreased from 4.67 dB for 3-min time scale to 1.56 dB for the daily scale ".

OK, so much for the abstract. I hope it is clear that you need to rewrite it.
The abstract was rewritten:
"Runoff and flash flood generation are very sensitive to rainfall’s spatial and temporal variability. The increasing use of radar and satellite data in hydrological applications, due to the sparse distribution of rain gauges over most catchments worldwide, requires furthering our knowledge of the uncertainties of these data. In 2011, a new super-dense network of rain gauges containing 13 stations, each with two side-by-side gauges, was installed within a 4 km^2 study area near Kibbutz Galed in northern Israel. This network was established for a detailed exploration of the uncertainties and errors regarding rainfall variability within a common pixel size of data obtained from remote sensing systems for time scales of 1-min to daily. In this paper, we present the analysis of the first year’s record collected from this network and from the Shacham weather radar, located 63 km from the study area. The gauge–rainfall spatial correlation and uncertainty were examined along with the estimated radar error. The nugget parameter of the inter-gauge rainfall correlations was high (0.92 on the 1-min scale) and increased as the time scale increased. The variance reduction factor (VRF), representing the uncertainty from averaging a number of rain stations per pixel, ranged from 1.6% for the 1-min time scale to 0.07% for the daily scale. It was also found that at least four uniformly distributed rain stations are needed to adequately represent the rainfall (VRF<5%) on a typical radar pixel scale. The difference between radar and rain gauge rainfall was mainly attributed to radar estimation errors while the gauge sampling error contributed up to 22% to the total difference. The ratio of radar rainfall to gauge-areal-averaged rainfall, expressed by the error distribution scatter parameter, decreased from 4.67 dB for 3-min time scale to 1.56 dB for the daily scale. The analysis of the radar errors and uncertainties suggest that a temporal scale of at least 10-min should be used for hydrological applications of the radar data. Rainfall measurements collected with this dense rain gauge network will be used for further examination of small-scale rainfall’s spatial and temporal variability in the coming years ".

Introduction

Line 26. Replace “ground truth” with “ground reference.”
Replaced, as suggested.

Page 5, line 1. I think you meant . . .correlation function not correlation coefficient

We believe that "correlation coefficient" accurately describes the referenced paper as most of Habib et al. study focuses on this subject.

Data

Line 7. The gauges are not "coupled", I think that a better and simpler way is to talk about the number of station, each with two side-by-side gauges. Or, double-gauge station. . .

As the reviewer suggested, we have rewritten line 7: "The rain gauges are deployed in 14 stations **each with two side-by-side gauges** (as in …". Also in page 11, line 3 and page 9, line 24.

Line 15. I do not think that Villarini et al. (2008) should get the credit here. Instead, you should cite those who conceived, designed and deployed the network.

Right, Villarini et al. used this network later, the credit should go to Moore et al. (2000). Reference corrected as follows:

Moore, R. J., Jones, D. A., Cox, D. R., and Isham, V. S.: Design of the HYREX raingauge network, Hydrol. Earth Syst. Sci., 4, 521-530, doi:10.5194/hess-4-521-2000, 2000.

Line 25. What pulse? Just say "maximum sampling frequency of 1 Hz". BTW, in the reminder of the section you never tell us how you collected tip data. Did you use the maximum sampling frequency to essentially record time-of-tip or did you accumulate (count) the number of tips in a given interval (I would guess 1 minute). As Ciach (2003) has clearly demonstrated, interpolating between tips results in more accurate estimates of rainfall amounts, especially at shorter time scales and lower rainfall intensities.

Line 25 changed as the reviewer advised to: "The maximum sampling frequency of the data logger is 1 Hz, …".

We accumulated the number of tips in 1-min time intervals (as the reviewer guessed) and then interpolated the data as described in lines 11-15 on page 7. Information was added to line 4: "The accumulated rainfall for this period**, collected by counting the number of tips in a 1-min time intervals,** is equal to 512 mm …".

Page 8, line 7: mean elevation? What do you need the "mean" for? Is this the elevation of the radar beam over the network? The last sentence in this paragraph is awkward too. What is "substantial" ground clutter? BTW, have you ever detected effects of AP in those 12 pixels?

Mean elevation- it is the elevation of the radar beam. We changed line 7 to: "(elevation of 710 m above **the network**).

"substantial"- we meant to say there weren't any constant ground clutters is these pixels.

AP effects have not been detected for the study region.

Next paragraph. Please clarify that saying "the same period" accounts for the days (hours) of radar malfunction and maintenance.

Yes, it is the same. We deleted "the same period" from this sentence to avoid confusion.

The next paragraph (line 16-24) is the most troubling for me. Why would you derive a separate Z-R for each pixel? First, on what basis, and second, what for? I understand the need to eliminate the overall bias from the radar data. The bias due to radar electronics is the same for all the pixels. Also, all pixels experience roughly the same storms. Therefore, you should adjust all pixels for the same bias and live with the consequences.

Separating the systematic from random effects is one of the fundamental difficulties in rainfall remote sensing. They affect each other but there is little you can do about it. Your estimate of the bias (one value) is just that: an estimate. Also, by studying the effects of a single Z-R you would make the study more relevant to operational applications.
Since our main objective in this study is to examine the rainfall distribution within a typical radar pixel we wished to compare the gauges over the study area to a single pixel, but the study area is overlapped by more than one radar pixel (see Figure 1). Therefore, we decided to make the Z-R adjustment for each radar pixel separately, as if it entirely covers the rain gauge network.

In the same paragraph. Is the upper threshold the only way to eliminate the hail cases? What about the bright band effects? What do the gauges show for the alleged hail cases?
Last year when we collected the data, the upper threshold was the only way to eliminate hail cases. From this winter we add a disdrometer to the study area, so hail that reaches the ground can be also detected. In this paper we didn't examine how hail affects the rain gauges.

Section 3

Line 16. It is more conventional to call nugget (1-c1).
We changed the order of the sentence, as suggested: "where *c*1 represents the **nugget (**zero-distance correlation**)** …".

Page 10, top. In fact, most networks used double-gauge stations.
Indeed, lines 2-3 were deleted.

Page 10, line 6-11. The fact that the network(s) cover only limited range of distances is only one potential reason for the existing discrepancy between the results reported in the literature. Others include sample size differences, estimation methods, statistical artifacts (e.g. bias in the correlation coefficient due to the skewness of the rainfall distribution).
We agree. We have added the reviewer remark in line 9: "**The disagreement between the correlation distances can be explained by sample size differences, different estimation methods or statistical artifacts. In addition, as** the spatial scale …".

Section 4
Page 10. The VRF is due Mejia and Rodriguez-Iturbe (you can just cite the text by Bras and Rodriguez-Iturbe or the references therein). Morissey et al. (1995) proposed a numerical method for calculating it. The method allows accounting for many different covariance models and arbitrary configurations of the investigated network. Krajewski et al. (2000) investigated in detail different aspects of the method application.
Line 24 was changed as follows to give the VRF its right credit: "The VRF methodology was introduced by **Rodríguez-Iturbe and Mejía (1974) and Bras and Rodríguez-Iturbe (1976).** Morrissey et al. (1995) **proposed a numerical method which** considered the number of rain gauges, their spatial distribution, and the correlation between them".
Referenced added as follow:
Rodríguez-Iturbe, I., and Mejía, J. M.: The design of rainfall networks in time and space, Water Resour. Res., 10, 713-728, 10.1029/WR010i004p00713, 1974.

Bras, R. L., and Rodríguez-Iturbe, I.: Evaluation of mean square error involved in approximating the areal average of a rainfall event by a discrete summation, Water Resour. Res., 12, 181-184, 10.1029/WR012i002p00181, 1976.

Page 11, bottom paragraph. When you talk about how small the VRF is, note that its standard deviation equivalent is much larger.
Right, but here we try to find only the best configuration, thus we looked only for the minimum VRF values.

Summary

I think that the authors should rewrite this section in the spirit of the title instead of merely repeating the just-presented results. What are the lessons learned? What else do we need to learn?
We agree. This was also mentioned by the second reviewer. We decided to change section 6 from "Summary" to "Conclusions", where we also discuss the lessons learned from the first year of observation. Here is the suggested re-written section:
"Subpixel rain distribution was investigated using a high-density network of rain gauges within a 4 km^2 area as a part of continuous efforts to better understand the uncertainties and errors of rainfall estimation at this scale. This is of particular importance when using remote sensing rainfall data (from ground weather radar or from satellite) for hydrological applications. In this study, we used the network of 27 tipping-bucket rain gauges located in northern Israel to evaluate the Shacham weather radar's performance and to learn about small scale rainfall variability. From the first year of observation three lessons were learned:

First, we examined the spatial correlation of gauge rainfall data as has been done before for different locations worldwide. We found that the nugget (zero-distance correlation) between rain gauges is high (0.92 for the 1-min time scale) and it increases with increasing time scale. Moreover, spatial rainfall correlations for all separation distances generally increase with time scale. The more important finding was that there is a difference in spatial correlations between convective and nonconvective rainfall, where the convective rainfall correlation decreases much faster with distance than the nonconvective. Convective rainfall correlations have long been a subject for study (for example, see papers by Sharon, D., 1974 and Osborn et al., 1979) but, to the best of our knowledge, this is the first time that this examination has been performed at the radar pixel scale. The fast decay of convective rainfall correlation within a radar pixel may imply that the radar errors for high rainfall intensity are even larger than thought. Further investigation is needed to understand the spatial and temporal differences between the different types of rainfall and its effects on radar data.

The second lesson learned was derived from the variance reduction factor. It was found that the VRF decreases as the time scale increases, from 1.6% for the 1-min scale to 0.07% for the daily time scale. This led to the conclusion that for any given time scale the average rainfall derived from the gauge network well represents the pixel-scale. This raises the question of how many rain stations are needed within a radar pixel for a good representation of rainfall at this scale. We found this question important for regular maintenance of the network (for example, when it is necessary to remove some rain gauges for calibration) and for future planning of other networks in similar climatological conditions (as the VRF is dependent on rainfall correlation). If the 5% threshold is selected as a criterion for adequate representation of subpixel rainfall distribution then, according to our analysis, four uniformly distributed rain stations are sufficient to represent the rainfall within the radar pixel. At least eight rain stations are required to represent the radar rainfall with a VRF threshold of 2%. The decision as to which VRF threshold to use however remains subjective. This finding can be used in validation procedures of remote sensing rainfall products with a similar pixel size. In the majority of cases, only one rain gauge is located within each validated pixel, while our results indicate that to remove uncertainties related to subpixel rainfall distribution four rain gauges per pixel are better, at least for a similar climate. Obviously, this will increase the cost of such a network; however, it will also assure meaningful validation results.

Lastly, we debate the question of which radar temporal resolution should be used for hydrological applications. A possible answer to this question is to fit the required temporal resolution to the basin hydrological response that depends on catchment size, land use and other properties (Morin et al., 2001). Berne et al. (2004) suggested a temporal scale of 3-5 min for urban catchments of the order of 1-10 km^2 while Atencia et al. (2011) suggested a temporal scale of 12-15 min for basins of the order of 100-1000 km^2. Another aspect to consider is radar errors for the pixel scale and its change with time (Fig. 7). It was found that the radar–rain gauge error decreases from 486% for the 3-min time scale to 344% for the 10-min scale and down to 56% for the daily time scale (all area mean values). This error is mainly the result of radar estimation errors, as the gauge sampling error contributes only 12%–22% to the total error, depending on the time scale. The improvement in radar rainfall estimations with increasing time scale is reflected by the increase of the CSI and POD parameters with time scale and the simultaneous decrease of the FAR parameter. In addition, the radar-to-true rainfall ratio, expressed by the scatter parameter, decreases with increasing time scale from 4.67 dB for the 3-min scale to 3.83 dB for the 30-min scale and onward. Based on these results, we recommend utilizing the radar rainfall data at scales of at least 10-min, thus benefiting from the large reduction in error from 3-min time scale to the 10-min scale.

We intend to continue collecting rainfall measurements with this network of rain gauges in the years to come. In December 2012, a disdrometer was installed at this site to measure rain drop size distribution (e.g., Jaffrain et al., 2011). We are looking for new and better ways to continue developing this network for future use with other weather radar or satellite observations".

Referenced added:

Sharon, D.: The spatial pattern of convective rainfall in Sukumaland, Tanzani — A statistical analysis, Arch. Met. Geoph. Biokl. B., 22, 201-218, 10.1007/bf02243468, 1974.

Osborn, H. B., Renard, K. G., and Simanton, J. R.: Dense networks to measure convective rainfall in the southwestern United States, Water Resour. Res., 15, 1701-1711, 10.1029/WR015i006p01701, 1979.

Atencia, A., Mediero, L., Llasat, M. C., and Garrote, L.: Effect of radar rainfall time resolution on the predictive capability of a distributed hydrologic model, Hydrol. Earth Syst. Sci., 15, 3809-3827, 10.5194/hess-15-3809-2011, 2011.

Morin, E., Enzel, Y., Shamir, U., and Garti, R.: The characteristic time scale for basin hydrological response using radar data, Journal of Hydrology, 252, 85-99, 10.1016/s0022-1694(01)00451-6, 2001.

Figures

Figure 1. Looks good. I would remove the lat/lon coordinates on both the inset and the main panel (nobody will use those for navigation. . .). Also, you should indicate the direction towards the radar.

We have removed the lat / lon on the major map and added a direction arrow toward the radar, as suggested.

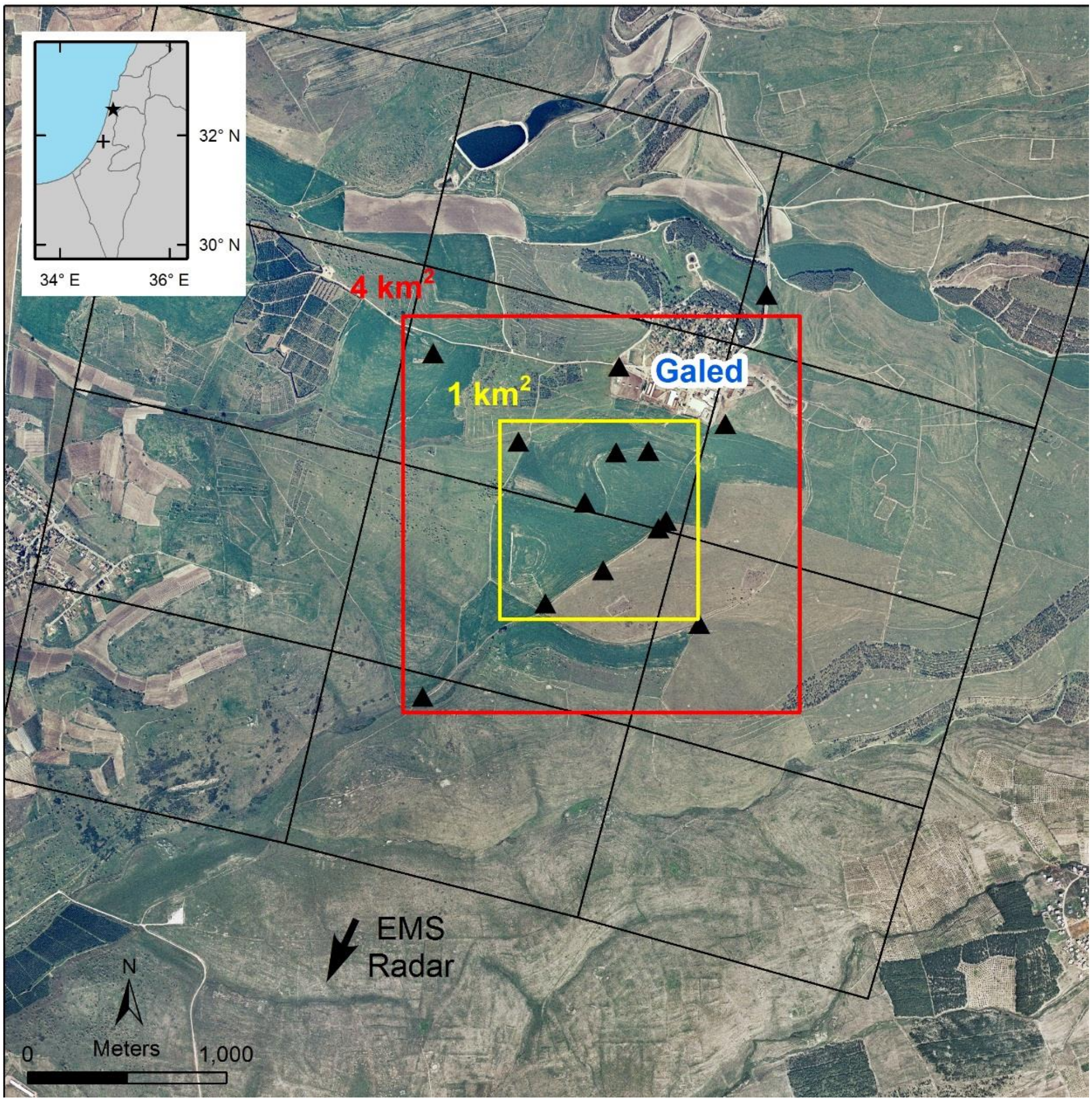


Figure 2. I think that you should use the same scale for all three panels, remove the scale description between the panels to make them larger. You call these observations "synchronous." This implies that your clocks are synchronized. But the HOBO loggers are notorious for having clocks drift, sometime substantially. How did you assure time agreement? If you did not, you should point out that this is another source of the scatter.

The scale of the panels was changed and the figure enlarged as suggested.

The HOBO loggers were synchronized each time the data were downloaded and checked during the QC for clock drifts- when problems found the data were not used.

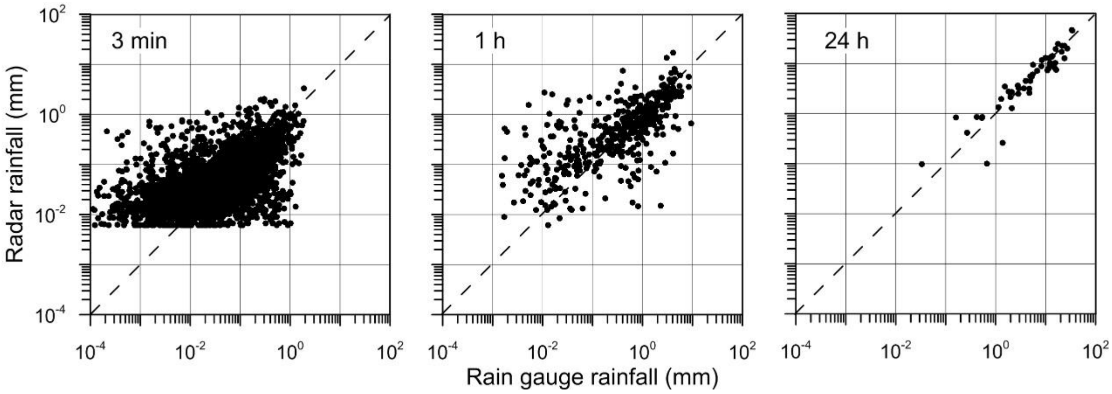


Figure 3. I do not understand what the points are. Please explain with a formula. Judging from the Table 1, the distances between points should not be evenly spaced but looking at the plot it seems that they are. Also, it is amazing that there is almost no scatter. How did you handle the zeros? What about the bias in the correlation coefficient? It may be better to just list the values of the parameters of the correlation function. Writing them in the equation form makes it difficult to read (symbols are too small).

The points represent the Pearson correlation between all pairs of rain gauges that are in the range of distance (lag) h of each other, as defined in equation (1). The zeros were also considered in the correlation analysis. The small scatter is quantified well by the three-parameter exponential function presented. The zero-distance is calculated from Pearson correlation of each two side-by-side rain gauges (each station), as explained in page 9, lines 21-24. The parameters of the correlation function are also presented separately in Figure 4.

Figure 4. This figure is begging for a vertical arrangement of the panels. . . So is Figure 8.

Right! Both figures (4 and 8) were re-arranged.

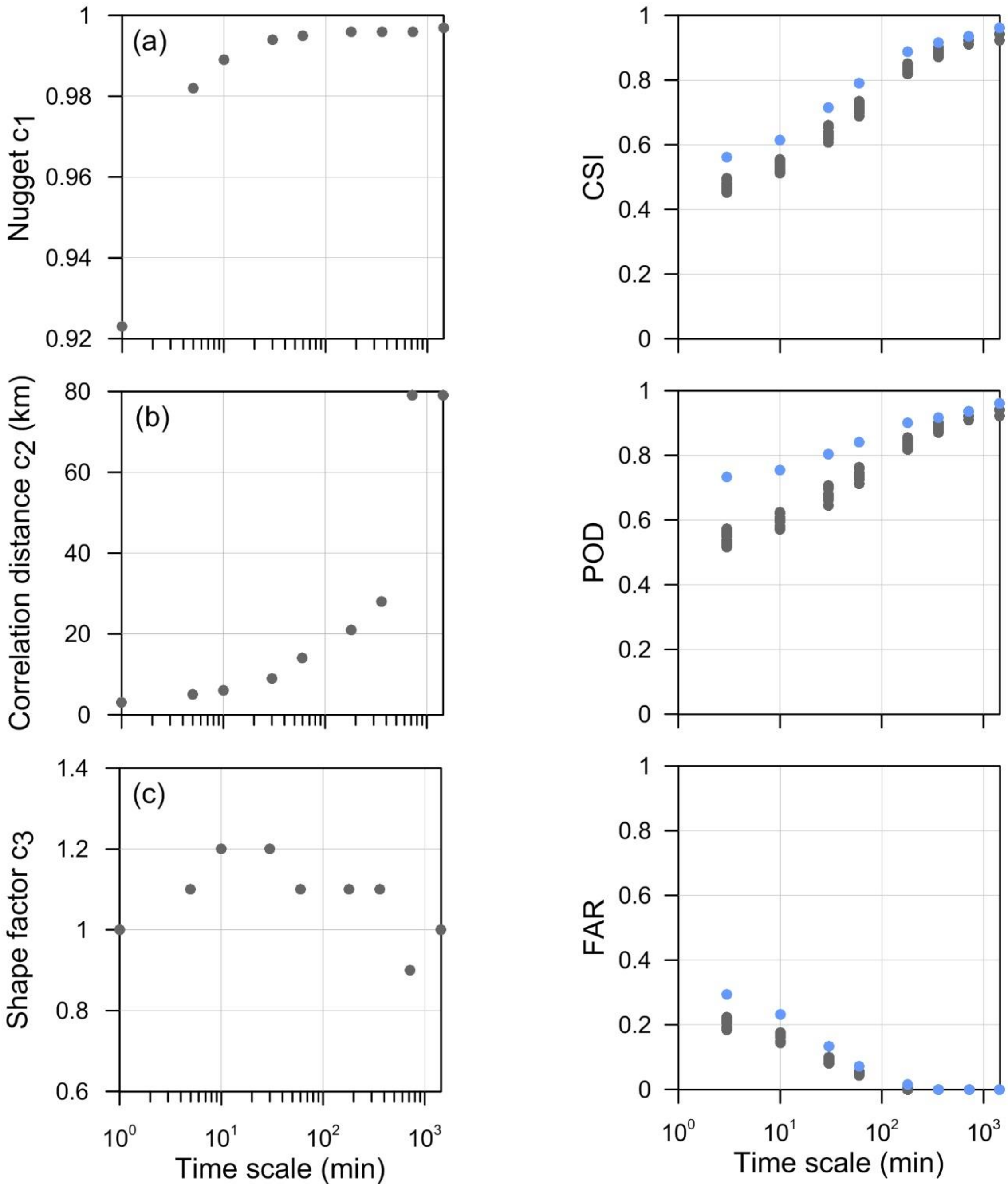


Figure 5. The caption should provide more details so that the figure is self-described. Also, in panel (b) it seems that the authors report only averages. There should be a scatter associated with each number of gauges (as there are many combinations of 2 out of 13, for example).

We prefer to keep the caption as short as possible, directing the reader to the manuscript for further information about the VRF.

In panel (b) the scatter represents only the best (lowest) VRF results from all possible combinations, as explained on page 12, lines 4-6.

Figure 6. Similar puzzle as for Figure 3. The plot (a) indicates that the authors have a gauge (station) pair every 200 m but Table 1 says otherwise. Also, the vertical axis

description is too dense. Since you have a grid you do not need labels every 0.1 for the correlation.

Plot (a) represents the correlogram h separation distance and not the distance between the gauge stations; see detailed answer regarding Figure 3.

We removed a part of the vertical axis description to make it less dense.

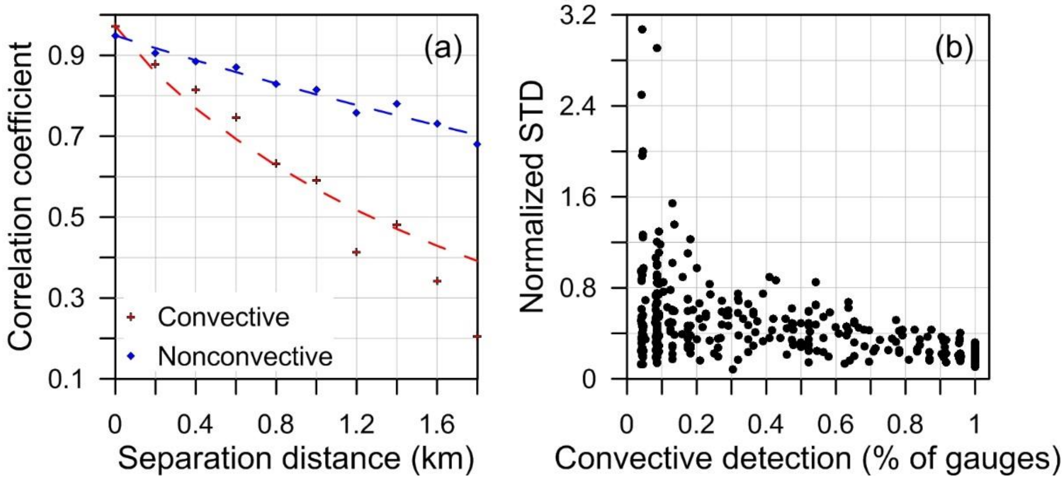


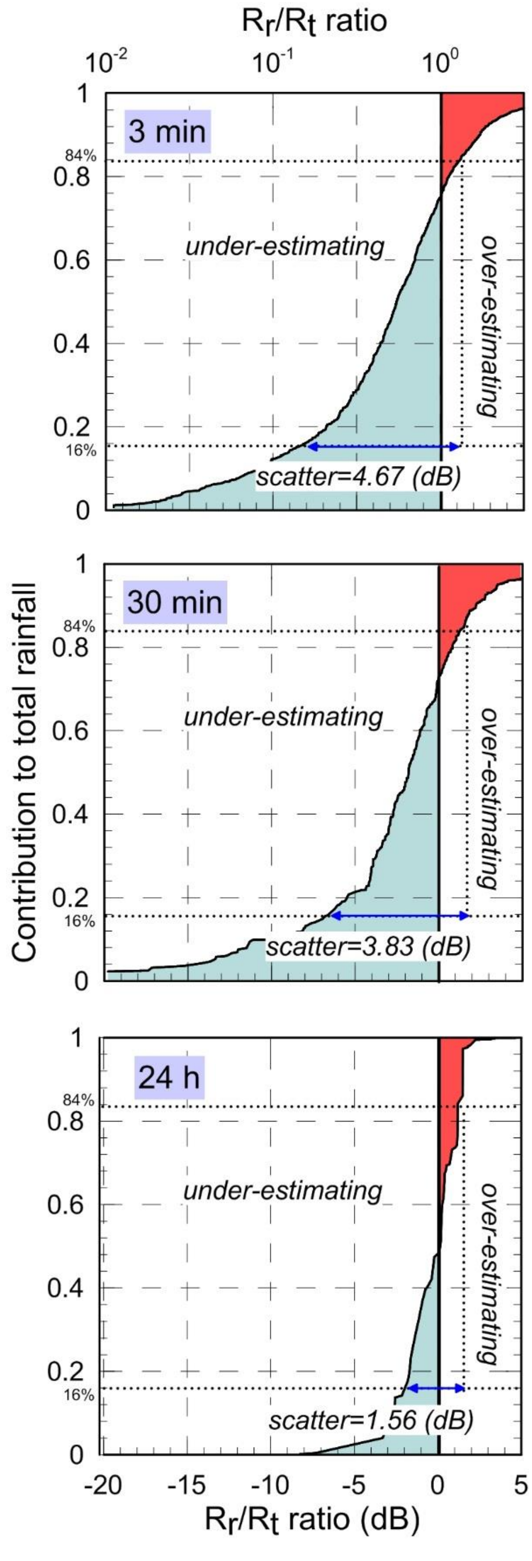
Figure 7. I am totally confused. Please explain precisely what the figure is supposed to show. My guess is that it is supposed to show the contribution of the radar-rainfall error and the rain gauge representativeness error to the variance of the difference between the two. But the caption says (at the end): "true rainfall derived from 12 radar pixels." True rainfall from radar??? I'm lost. Also, please provide details with respect to scales in space and time. I was under an impression that you computed the VRF for the 4 km^2 area but it seems that if you are comparing the average of the 12 radar pixels you should calculate the VRF for that area as well.

You guessed right, we probably explained it poorly. First, the caption was corrected as it should be: "The normalized root mean square errors of radar rainfall (R_r) vs. true rainfall (R_t) were analyzed independently for 12 radar pixels. The maximum and minimum are presented by the blue sections". Of course, the true rainfall is unknown (therefore we used equations (4)-(6) to bypass this unknown value) and because of the large error of the radar we can't even assume that we know the true rainfall.

The VRF was indeed computed for the 4 km^2 area. This area is overlapped by at least 4 radar pixels (see Figure 1) and perhaps also by the 4 leftmost pixels, as the angle of the radar may not start exactly from the North. For this reason we examined all the 12 radar pixels near the study area. Please note that the area of the radar pixel is about 1.6 km^2. Although we are not making a direct comparison between the results of the VRF and the results of the ESM in the paper, we claim that the 4 km^2 area represents the same scale of a one radar pixel.

Figure 9. The presentation in Figure 9 (following German et al. 2006) is interesting but the figure is too small. It too would benefit from a vertical arrangement and an increased size.

Done!



Reviewr #2

General Comments

This paper describes a study of small-scale rainfall variation and its effect on estimating area-average rainfall from rain gauges, that is then used to determine the uncertainties in radar rainfall estimation. The paper is well-referenced and well-written. However, the authors need to make very clear what the novel contribution of this paper is. There are several issues with the analyses that need to be addressed, particularly those related to the sampling of tipping bucket rain gauges at short time scales. Finally, I think that the authors should devote more time to drawing conclusions from the results (instead of just presenting a summary in the last section. I have several specific comments that are given below.

Specific comments:

On p. 2, line 11, I suggest noting here that the distance to the radar is 63 km.

It has been added to the abstract, as suggested.

On p. 2, line 20, the authors note that the "radar-to-true rainfall ratio decreased with increasing time scale". The fact that the ratio decreases could mean an improvement in case of radar overestimation, but can be a degradation in case of radar underestimation. I think the authors should make clear whether this is an improvement.

True, it has been corrected to: "The radar rainfall estimations were improved with increasing time scale and the ratio of radar rainfall to gauge-areal-averaged rainfall, expressed by the error distribution scatter parameter, decreased from 4.67 dB for 3-min time scale to 1.56 dB for the daily scale".

In the Introduction, I think the authors should make very clear here what the novel contribution of this paper is. What sets this paper apart from the literature that is cited in this section?

As the reviewer suggested, we decided to broaden the last paragraph in the introduction as follows:

"In the current study we set up the first step toward estimating the subpixel sampling uncertainties and the errors of weather radar rainfall estimates using a super dense rain-gauge network. This network is **located in a different climatological area than the above presented networks. In this paper we wish to: first, present our network (current results and future plans) to the hydrological community as part of the global effort to enhance the knowledge of small scale rainfall variability and radar uncertainty; and secondly, to present three lessons learned from the first year of observations regarding (1) the spatial correlation of convective and nonconvective rainfall; (2) the number of rain gauges required to adequately measure rainfall in a radar subpixel scale; and (3) the decision as to which radar temporal resolution should be used for hydrological modeling due to the radar errors at the pixel scale.** The paper is composed of **six** sections … The conclusions**, discussion on the lessons learned,** and near-future plans for the rain-gauge network are presented in Sect. 6".

In the first paragraph of p. 3, the authors could consider mentioning the importance of spatial rainfall variation for hydrology in urban areas (see e.g. Berne et al., 2004, Journal of Hydrology 299, 166-179 or Smith et al., 2002, Journal of Hydrometeorology 3, 267-282).

This paragraph was extended to include Berne et al. paper, as follows: "**; Berne et al. (2004) found that hydrological applications for urban catchments of the order of few square km require high resolution of temporal (3-5 min) and spatial (2-3 km) rainfall data".** Reference added as follows:

Berne, A., Delrieu, G., Creutin, J. D., and Obled, C.: Temporal and spatial resolution of rainfall measurements required for urban hydrology, Journal of Hydrology, 299, 166-179, 10.1016/j.jhydrol.2004.08.002, 2004.

On p. 7, lines 16-25 the authors describe the QC procedures that were applied to the rain gauge data. It is not clear to me how this quality control is carried out. What were the criteria to remove data? Was this done by hand, or were there objective criteria?

First we plotted the accumulated rainfall of each rain event for each couple of side-by-side gauges. Times where one of the gauges measured much higher rain intensity than the other (using a threshold selected subjectively) were removed from this point to the end of the event. We then repeated the same procedure but for all gauges for each event and removed problematic sections. No objective criteria were used.

On p. 8, line 7, the authors indicate that the mean radar beam height at the location of the gauges is 710 m. What are typical wind speeds and directions occurring in rainfall events? This is important because rain may be blown away from directly under the respective radar pixel. Another issue is that there is some time between measurement by the radar and the arrival of the raindrops at the gauge (on the ground). Especially at short time scales this can have a significant effect (Leijnse et al., 2010, Journal of Hydrometeorology 11, 1322-1329).

We decided to compare the rainfall from the gauges to 12 radar pixels surrounding the study area (Fig. 1) as we can't be sure which of the radar pixels best represent the study area. This is because, as the reviewer mentioned, there are winds (mainly westerly) that may shift the rain (a previous study we conducted shows a mean velocity of about 11.8 m s^{-1}). We haven't dealt at all with the time lag between the radar records and actual arrival time on the ground – we believe it is an important issue which we plan to investigate in the future.

On p. 8, lines 16-24, the authors state that they use different Z - R relations for all radar pixels under consideration, each of which has been optimized using gauge data. This results in a huge (factor of 2.4) variation in optimal Z - R relations within a 4 km X 4 km area. But even if this variation would have been limited, the use of different Z -R relations over such a small area does not make sense to me. The fact that there are such great differences should be thoroughly investigated. It seems to be related to the first two digits (possibly azimuth; could this be partial beam blockage?) in the radar ID given in Table 2 (a ~ 60 for ID=10xxx, a ~ 90 for ID=11xxx, and a ~ 125 for ID=12xxx). In any case, I find the use of such wildly varying Z - R relations over such a small area unacceptable.

We must agree with the reviewer – usually all radar pixels are adjusted using one Z – R relation for the domain. But in this study our goal was not to have good rainfall estimation for the whole 12 radar pixels area, as in a standard gauge adjustment procedure, but rather to focus only on the 4 km^2 area where our network is located. To resolve the uncertainty (i.e., which of the radar pixels represent the network area) we repeated the analysis 12 times; each time another pixel was assumed to be above the network and therefore was adjusted with its gauges and resulted in a different adjustment factor. The reason for the large difference we got with azimuth (as the reviewer points out) could be related to local rainfall gradient (increasing from west to east), that could result in an increase of the 'a' parameter in the eastern direction. However, there is not enough data to test this.

On p. 9, line 2, the authors refer to "the problem mentioned in the previous section". I think this problem should be stated here explicitly.

We agree, that sentence was modified to: "… possibly due to the backward linear interpolation that was conducted to overcome the unknown tipping-bucket fill time (see previous section)".

On p. 9, lines 19-21, the results of the correlation analyses are presented. I think that, especially for the shorter time scales considered here, the fact that the rain gauges are tipping bucket gauges can play a significant role. For example, on a 1-minute time scale, the minimum rainfall intensity recorded by a rain gauge is 6 mm h^{-1}, given a tip volume of 0.1 mm. And there will often be minutes with zero rainfall if the intensity is less than 6 mm h^{-1}. This issue should be discussed here.

The minimum rainfall intensity could be lower than the 6 mm hr^{-1} intensity that was actually recorded, as a backward linear interpolation was applied (see Section 2.1).

On p. 9, line 21, I think there is a typing mistake: "toward the daily scale" should be "toward 1 on the daily scale".

It does sound better, this sentence was corrected as suggested.

On p. 9-10, Section 3, the authors compare parameters of fitted correlation functions to those presented in the literature. I think that the authors should explicitly note here that these parameters need not be similar to those presented in literature because of different rainfall climatologies.

At the end of the section (p. 10, line 19) it is mentioned that differences are expected between experimental studies also due to different precipitation type.

On p. 10-12, Section 4.1, the authors discuss the variance reduction factor. I agree with them that a detailed discussion of the VRF is not necessary here, but the reader should be able to understand what VRF means. For this it is essential to clearly define both $\sigma^2_{\overline{R_s}}$ and $\sigma^2_{R_s}$. What is meant by "variance" here? For example, is $\sigma^2_{R_s}$ the uncertainty in gauge measurements, the spatial variation in point rainfall within a given area, or something else?

To clarify the explanation, we changed Eq. (2) and lines 3-7 as follows:

"Let R_s be the point rainfall of a single rain station (two side-by-side gauges per station), $\overline{R}_s$ be the averaged-areal rainfall derived from the rain stations, and let R_t be the true areal rainfall. The mean square error of the true rainfall to the averaged-areal rainfall can be expressed as:

$$E\left[\left(R_t - \overline{R}_s\right)^2\right] = \sigma^2_{R_s} \cdot VRF$$ "

We decided not to enlarge further on the method, as it has already been discussed in detail by Krajewski et al. (2000) and Villarini et al. (2008).

On p. 11, line 14, I think there is a typing mistake: "and is a Boolean" should be "and $\delta(i, j)$ is a Boolean".

Right! The delta was missing, it has been corrected.

I don't understand the sentence starting on p. 11, line 24 ("The VRF is very close to..."). Why does the fact that the VRF is close to zero imply that the mean square of the point variance is also close to zero? And what is meant by the "mean square of the point variance"? How should I interpret this?

The text was corrected to: "The VRF is very close to zero, meaning that the left side of Eq. (2) is also close to zero; thus for any given time scale, the true rainfall will be well represented by the **averaged-areal rainfall**".

On p. 13, line 5 the authors say that a percentage is plotted in Fig. 6b, and the label of the x-axis of that graph also reads “percentage”. However, what is plotted are fractions (i.e. a factor 100 smaller values). I suggest either changing the x-axis of Fig. 6b or changing the wording from “percentage” to “fraction”.
We changed the word “percentage” to “fraction”, as suggested.

On p. 14, lines 9-10, the assumption “we assume that there is no bias between the rainfall measured by the rain gauges and the rainfall measured by the weather radar” is valid, but not necessary if $\mathrm{var}\{R_r - R_g\}$ is not computed based on Eq. (4) but by just taking the variance of the differences between R_r and R_g.
The expression $\mathrm{var}\{R_r - R_g\}$ was computed based on Eq. (4).

On p. 14, there is a typing error: “reduce” should be “reduced”.
We have corrected the typing error.

On p. 15, the statistics CSI, POD, and FAR for the detection of precipitation by radar are evaluated. These statistics tell us something about the ability of the radar for determining the difference between dry weather and precipitation (no matter how intense). I don’t think these statistics are very relevant for hydrological applications.
We think that these statistics are relevant for hydrological applications as they help us to validate the general accuracy of the weather radar for this small study area for different time scales (this, in addition to the ESM presented in the previous section). Moreover, it is very important for hydrological purposes to detect the wet and dry periods correctly. These parameters can give us a general idea about the ability of the radar to detect these periods correctly. We believe that it is important to keep these statistics in the paper.

On p. 15, line 9, the authors mention that a “zero threshold was used to mark the occurrence of rain”. However, for both radar and rain gauges it is unclear what this means. For radar, there is always a signal (noise). I think that if the authors decide to keep the analysis of these statistics, it is essential to discuss how “zero” is defined for radar rainfall estimation (is this anything below the noise level?). For tipping-bucket rain gauges, “zero” rainfall is also ill-defined, so it should also be made very clear what is meant by “zero” rainfall for gauge measurements as well.
Agreed, for the reader to understand the evaluation we made, the definition for "zero" (both for the radar and the rain gauges) must be stated. We continued line 9 as follows: "A zero threshold was used to mark the occurrence of rain**. This means that the lower threshold for the radar was defined as 0.1 mm h^{-1}, while the averaged gauged rainfall was indicative for rain as at least one rain gauge recorded rain".**

On p. 15, line 19, the authors note that “an improvement was detected”, but I’m don’t understand what caused this improvement.
This "improvement" was the result of changing the radar’s lower threshold to a higher one. As the "improvement" was detected only for some, but not all pixels, we decided to remove this sentence.

On p. 15-16, the ratio of radar rainfall estimated to true rainfall is discussed. Were all radar pixels used separately, or were these averaged?
They were averaged, as mentioned on p.15, line 20.

On p. 16-17, Section 6, I think it would be better to have a section called "Conclusions", where conclusions are drawn about what we can learn from this study. I don't think the current Section 6 (Summary) adds much to the paper.
We agree, this was also mentioned by the first reviewer. We decided to change section 6 from "summary" to "conclusions" (as suggested), where we also discuss the lessons learned from the first year of observation. Please find the re-written section in the response submitted to the aforementioned reviewer.

On p. 24, Fig. 1, would it be possible to also show the location of the radar in the inset map of Israel?
Yes, the radar location was added (see figure 1 in our response to reviewer 1 above) and we also changed the figure caption to: "... Inset shows the general location of the network in Israel **(star) and the location of the weather radar (cross)**".