Responses to the reviewer comments
Paper: High resolution rainfall-runoff measurement setup for green roof experiments in a tropical environment
Authors: A.J.J. Vergroesen et al.
MS No.: hess-2010-349
MS Type: Research Article

General responses	
paper format	We tried to write the paper in a non traditional format that is promoted by Simon Peyton Jones (http://research.microsoft.com/en-us/um/people/simonpj/papers/giving-a-talk/giving-a-talk.htm). That implies that: • The literature review (traditionally in the introduction) is moved to the discussion at the end of the paper. • First the general idea is given. The details are presented later. This leads to some inevitable repetitions. Obviously we did not succeed to convince the reviewers. Therefore we will rewrite the paper in a more traditional format by moving the literature review to the introduction, and by combining chapters 2 and 3, thus removing the repetitions.
focus	New in our approach is the way we measure the runoff. A combination of a tipping bucket flow gauge and a high precision weighing scale results in continuous runoff information that will provide insight in the dynamic of the runoff. This provides us a better possibility to try to determine (model) the relevant runoff processes in green roofs. Next to that we try to show that the retention performance of green roofs should not be compared to rainfall (like it is traditionally done), but to a reference roof. Despite calibration and relative high resolution, rainfall measurements often result in underestimation of rainfall.
other aspects	We performed some evaporation measurements. The high resolution measurement results are used to determine the soil characteristics of the growing medium. The bathroom scale measurements are merely to (roughly) verify the conclusions on evapotranspiration that are drawn from the rainfall-runoff measurements, and to provide a rough estimate of the stored water volume. That is the reason we have taken these measurements up into the paper.
budget	The budget for the experiments was limited. Therefore we focused on the runoff measurements. We had not sufficient means left to install a wind insensitive rain gauge, let alone a high precision weighing scale that can carry a setup larger than 1 m^2 and heavier than 200 kg.
larger scope	The measurement method is indeed part of larger thesis. We decided to split this part from the measurement results because of the size of a combined paper, and because we want to elaborate on the novelty of the way we measure the runoff as well as the possible errors this way of measuring can introduce.
preliminary versions	Rewriting the manuscript is a major effort that will take considerably more time than is available to respond to the reviewers. However, to give the reviewers an impression of the rewritten version we added

	preliminary versions of abstract, introduction and rain gauge calibration at the end of this document.

Answers to individual reviewer comments:

G. Kulkarni	
comment 1	Are there any other research groups who have performed this experimental approach in the past. I did not see any references in the introduction section.
answer 1	We could not find any references of the approach we used to measure the runoff. We will reformat the paper, and move the literature review to the introduction section (see general responses, paper format)
comment 2	Using a weigh scale from Ikea, shown in Fig 4. It is known that these scales suffer from hysteresis when used under continuous load. That means there is a shift in the observed values with time under constant mass loading. Wondering if you ever observed this weigh scale behavior.
answer 2	We are well aware of the large inaccuracy of this way of weighing. It was carried out during a relatively short period, to give us a general idea of the storage / evapotranspiration in-between the ends of two consecutive runoff periods (Figure 9). We checked all scales before every measurement by adding the weight of a small person (approximately 50 kg). However this is no guarantee that there has not been a shift in observed values.

anonymous reviewer 1	
general comments:	
comment A	After reading the manuscript, it is not clear what the novelty of the work is. The loss of accuracy of tipping buckets for intense rainfall events is well established, the discussion on the calculation of runoff, although very detailed, is simply based on weighting water volumes as well as the estimation of evapotranspiration. The section on environmental disturbances on the measurements (section 4) is very general, and the discussion (section 5) looks more like a review of some other experimental setups with aims similar to those of the authors.
answer A	The novelty of our work is the way we measure the runoff. A combination of a tipping bucket flow gauge and a high precision weighing scale results in continuous runoff information that will provide insight in the dynamic of the runoff. This provides us a better possibility to try to determine (model) the relevant runoff processes in green roofs. Next to that we try to show that the retention performance of green roofs should not be compared to rainfall (like it is traditionally done), but to a reference roof. We are aware of the fact that the loss of tipping buckets for intense rainfall is well established. We merely want to point out that in spite all our efforts the reliability of our rainfall measurements is a lot less than the reliability of our runoff measurements. Therefore in green roof experiments runoff from a green roof should not be compared to rainfall but to runoff from a reference roof, since runoff from both roofs is measured in a similar way. That may seem obvious and everybody may be well aware of errors in tipping bucket rain gauges, but literature review shows that in most experiments green roof runoff is compared to rainfall measured by a tipping bucket rain gauge, generally without taking into account the effect of wind and the effect of tipping loss during more intense rainfall. We will rewrite the paper in another format, move the literature review from the discussion section to the introduction, shorten the paper and focus more on the novelty and the point we want to make regarding green roof runoff comparison to rainfall (see general responses, paper format and focus).
comment B	When reading it, the manuscript appeared to me like a chapter of a report or a thesis dedicated to the description of the methods used for a research project, but the results from the research (data analysis and discussion of the data) are missing.
answer B	The manuscript is indeed part of larger thesis. We decided to split this part from the measurement results because of the size of a combined paper, and because we want to elaborate on the novelty of the way we measure the runoff as well as the possible errors this way of measuring can introduce. As such measurement results shown in this paper are only used to elaborate on the methods we used.

specific comments and technical corrections:	
comment 1	The paper appears quite long and some concepts are repeated more than once. For example, the calibration of the tipping bucket appears in section 2.1 and 3.1. In section 2.3, the size of the box is repeated twice, once at line 14 and again at line 27. At page 9374, line 25 repeats line 6.
answer 1	This is partly caused by the writing format we applied (see general response, paper format). We will rewrite the paper in a more traditional format, move the literature review to the introduction, remove all data treatment aspects from section 2 to section 3 and prevent unnecessary repetitions. The final paper will be about 1/3 shorter.
comment 2	Some statements, even though true, are redundant. For example, at the end of page 9371 and the beginning of page 9372, the authors say that they transform runoff volumes in equivalent mm to compare to rainfall. This is rather common in hydrology and I would remove it.
answer 2	The second part of section 2.2 will be removed and combined with the text in section 3. We agree with the reviewer that the sentence "Transforming the runoff to mm per catchment area makes it easier to compare the runoff to the rainfall, which is often measured in mm." is redundant and will remove it.
comment 3	Page 9371, line 27: 2.51 mm/m^2 should be simply 2.51 mm.
answer 3	We will change 2.51 mm/m^2 into 2.51 mm
comment 4	Page 9376, line 3: ...weights have been transformed...
answer 4	We meant it slightly different. We will change "weights have be transformed" into "weights have to be transformed"
comment 5	References: many of the references are from conferences and difficult to find.
answer 5	We can only agree with the reviewer on this. However we were able to find these conference references. If required we can provide them to the reviewer. This stresses the importance of this method paper in a well-recognized journal like HESS.
comment 6	Figures: many figures are not useful and informative. I would remove the figures with pictures of equipment that can be easily found in user manuals from companies (rain gauges, scales, etc.).
answer 6	Regarding the figures on equipment we can remove Figure 1 (rainfall measurement setup). To our opinion Figure 2 on the runoff measurement setup can not be removed. Anonymous reviewer 2 suggested to remove two of the three figures on weather influences (Figure 13 and 14), as well as tables 3 and 4. We agreed to remove Figure 14 and tables 3 and 4. We proposed to remove Figure 12 in stead of Figure 14 (see anonymous reviewer 2, specific comments, answer 28).

anonymous reviewer 2	
general comments:	
comment A	Towards the end of this section, the authors mention the use of meteorological data to estimate evaporation (method?), which is not presented at all in the manuscript.
answer A	We did not measure these data ourselves; we used the data from the NUS weather station. These data were used in a later stage to model the processes in the unsaturated zone of the green roof. Model: Hydrus-1D; method: Penman-Monteith This will be made clear in the text.
comment B	The authors "assumed" that the factory calibration of 0.2 mm per bucket is correct, WHAT? A volumetric verification (with pipette) is ALWAYS the first step needed, in both lab and field installation. This calibration is done by using the two calibration stop screws (one for each bucket).
answer B	Since the manufacturer delivered an 0.01 inch based tipping bucket first we were very alert on calibrating. We did not assume a correct calibration; we wanted to do it ourselves. The reviewer is correct that a calibration should be started with a volumetric verification. In fact we actually did. As a result of which we found that the difference between the two tipping volumes is a little over 10%. That implies that the calibrated conversion rule results in an error of a little over 5% for a single tip (page 9374 line 20 – 25). We will include our calibration approach in the text, and change the description of the rain gauge calibration approach accordingly. A preliminary version is added at the end of this document.
comment C	There are other issues (which the authors are not aware of) associated with the setting up of the tip-counting scheme for a particular data logger. This error is of particular interest when the rainfall intensity increases. Some data loggers experienced recording problems when the time between two consecutive tips is less than 3 sec. The problem may results in either recording an extra count or not recording the tip at all, thus affecting the NRT.
answer C	The aspect of tip-counting was tested under several conditions: (1) During calibration the tips were counted simultaneously manually (tip sounds) as well as by the data logger. (2) After several events where runoff from the reference roof exceeded the rainfall manual tipping tests were conducted to check errors in tip-registration. The tip-frequency during these tests varied between 6 tips per minute to more than 100 tips per minute. The tests were conducted with a dry tipping bucket magnet, with a wet tipping bucket magnet, and even with a piece of paper that covered the magnet. Several thousand tips were all recorded by the data logger. The text will be adjusted to reflect this. A preliminary version is added at the end of this document.
comment D	The extreme event generated for calibration purposes in Figure 6 and its corresponding equation has an intensity of approximately 800 mm/h, resulting in an average tipping frequency of 67.5 tips/min or one tip

	every 0.8sec. This reviewer has two problems with this: what is the purpose of including an unrealistic intensity in a calibration equation for intensities values that exceed by eight times the operational limits by the rain gauge maker? and what is the error associated to NRT under this condition given the fact the time between two consecutive tips is very close to the "filling + half tip" time together?? It is hard to believe that half tip time alone issue will account for the increase of 60% in the average tip volume as presented in Figure 6.
answer D	The extreme test corresponds to an intensity of 22.5 mm/min (1350 mm/hr). It took the water more time to leave the funnel than it took to enter the funnel. Therefore the test resulted in an approximation of the funnel capacity. Since the test has been performed we presented it in the graph. Even though it is an unrealistic intensity the point is almost perfectly in line with the rest of the points in the graph. Using the half tip time of 0.41 sec that resulted from all tests and the frequency of the extreme test of 67.5 tips/min implies a total loss period of almost 28 sec each min. The resulted half tip time alone should increase the average tip volume by 85 %. The half tip time that resulted from this test alone is 0.34 sec, leading to an increase of 62 % of the average tip volume.
comment E	Even though the density of the liquid is a key player in the system (after all we weight the liquid as runoff), a constant density for the water has been assumed here. Did the authors measure the water temperature of the rainfall or runoff? How valid is this assumption? What is the impact of the assumption in the system performance as a whole? Water temperature has shown to introduce another source of error in using weight systems (see Sevruk and Chvíla, 2005, Error sources of precipitation measurements using electronic weight systems).
answer E	We did not take into account the temperature effects on the density of water for two reasons: (1) The collected water in the tipping bucket on the weighing scale is always in the shade. Therefore we assume that only air temperature can affect the water temperature. The air temperature in Singapore varies between 24 degrees centigrade during night time to 32 degrees centigrade during day time. That is rather constant. (2) We compare the green roof runoff to the runoff from the reference roof. The runoff from both roofs occurs almost simultaneously. Therefore we assume that the differences in water temperature between the two roofs are negligible. However the reviewer is correct that we did not mention this effect in the paper. The errors due to temperature are: 24 ^{o}C: 1000 g = 1.0026 liter → underestimation 0.0026 mm 32 ^{o}C: 1000 g = 1.0049 liter → underestimation 0.0049 mm We will include this in the text.
comment F	Something else is happening on the scale readings as a runoff event progressed, which may not be related to the tipping time and tipping action. Looking at Figure 3, it can be seen that as the rainfall event

	progresses the maximum weight of bucket with water (prior its tipping point) increases, it reaches a peak and then it reverses its trend. A similar behavior can be seen in Figure 15. It also seems to be happening for the low weight as the bucket empties. Initially this reviewer thought in a possible density effect as the cooler rainfall water will initially increase its temperature as it encounters a warm air and plate's surface (during the earlier stages of the event), but it is expected that it will keep its cooler temperature signature afterwards (heavier water) resulting in higher weight readings. But in 1 liter of water a 2 g difference is expected due to a change in density due to temperature (if nothing else like particles get into the bucket), which is much smaller than the 200 g difference (sometimes 400 g) observed from the records. Any comment on that?
answer F	During intense rainfall some water will can be temporarily stored in the funnel of the tipping bucket, increasing the total weight of the tipping bucket. The minimum weight increases as well. When runoff intensity decreases again the stored water in the funnel reduces to (almost) zero again. However we use the weight changes to determine the runoff and the runoff intensity (see Figure 7). Next to that the maximum recorded weight can be influenced by the exact moment of recording. The actual tipping generates relevant additional pressure (by impact) on the weighing scale. When the moment of recording is closely after the actual moment of tipping some additional weight will be recorded. This is explained in the paper (Figure 7, point 3A)
comment G	This reviewer suggestion is to work towards a better manuscript organization to improve readability and avoid unnecessary repetitions. Since source of errors for runoff and evaporation has been identified and quantified, these should be taken into consideration for runoff and evaporation estimates and properly discussed their effect on the final results. Evaporation estimates resulting from the different approaches need to be contrasted with evaporation calculation from the meteorological data available.
answer G	The method and error estimates are central to the innovation presented in the paper so this analysis will indeed be central. The change in organization of the paper will in this respect also be helpful.
Specific comments:	
comment 1	Page 9368, Line 4. Use "rainfall and runoff" or "rainfall-runoff" measurements.
answer 1	We will change "rainfall, runoff" into "rainfall-runoff"
comment 2	Keywords Suggest not using "high resolution measuring equipment" as a keyword.
answer 2	We will use "high resolution" instead
comment 3	Introduction: It does not introduce information on alternatives techniques available to measure runoff in small experimental green-roof experiments. Please, move the relevant literature from the discussion section.
answer 3	This is due to the paper format we used. As mentioned above in general responses, paper format we will rewrite the paper in the more traditional format.
comment	Page 9368. Line 14. It is not clear which hydrological processes the

4	authors refer to.
answer 4	We refer to processes that influence the runoff from green roofs, in the growing medium as well as in the drainage layer. That are processes like interception, percolation, evaporation, transpiration, hysteresis, storage, water flow through the drainage layer, etc.
comment 5	Page 9369. Line 22. Add (NUS) after National University of Singapore.
answer 5	We will add (NUS)
comment 6	Page 9369. Line 25: Authors presented a "discussion" about differences between rainfall measurements.
answer 6	We will remove this discussion here and simply state that we used a rain gauge on site and used evaporation related data from the NUS
comment 7	Page 9370. Line 15. The rain gauge needs an elementary volumetric calibration for each bucket prior to any other calibration procedure.
answer 7	Like mentioned in the general comments (answer B) we carried out such an elementary volumetric calibration. We will state more explicitly that we did.
comment 8	Page 9370. Line 24. An equation (result) is introduced. Move to the results section.
answer 8	We will move this to the data treatment section.
comment 9	Page 9371. Lines 20-25. Suggest moving Figure 3 and text below to "Data treatment".
answer 9	We will move this to the data treatment section.
comment 10	Page 9371. Line 20. The authors should assess the effect of water temperature.
answer 10	We will address the issue here, and elaborate on it in the data treatment section (see also general comments, answer E).
comment 11	Page 9372. Line 4. The authors have not mention vegetation type to account for transpiration due to lack of description on the green-roof experimental setup.
answer 11	That is correct. Although not relevant for the way of measuring we will add that the vegetation type is Sedum Gold Mound.
comment 12	Page 9372. Line 20. Reference to a soil water content storage. At this point the reader does not have information about this storage due to lack of description in the experimental setup.
answer 12	We will add a description of the experimental setup, referring to figure 2.
comment 13	Page 9373. Line 5. The example on evaporation calculation should be part of a Result section. Also, the evaporation value of 6.6 mm in calculated based on weight. This reviewer will come back later to this point.
answer 13	We will move this to the data treatment section.
comment 14	Page 9373. Line 20. Comment/suggestions: A real intensity data logger combined with a 0.1 mm bucket will help in reduce issues for low intensity rainfall events.
answer 14	We like to thank the reviewer for this suggestion. We have considered using a 0.1 mm bucket, but decided on a 0.2 mm bucket because of the often very intense rainfall in Singapore. Next to that we have been looking for a rain gauge data logger combination with a data logger that records the exact date and time of tips only. At the time

	of our search we did not succeed in finding one. However, as explained in the document (Page 9386 Line 5 – 10), a more exact tipping time for single tips than a single minute can create bogus accuracy.
comment 15	Page 9374. Line 4. ..”for 20 tests the total number of tips and the time of the last tip are recorded...” Is this information obtained from the data logger?
answer 15	No, this information is obtained by using a stop watch. The time of every tip is recorded. Only the last time is used in the calibration. The rest of the time recordings are only used to check whether the “rain-intensity” was somewhat constant. That was the fact. Of course there was some variation in inter-tip time, but this variation was limited. Besides that actual rainfall will probably also show some variation in intensity within a minute. We will add the fact that we used a stop watch for timing.
comment 16	Page 9374. Line 12: Replace “overview the calibration results” by “overview of the calibration results”. This reviewer will recommend rewriting this paragraph and delete reference to the extreme event calibration unless the authors provide evidences of no further potential issues with the NRT counts from the data logger (see general comments.)
answer 16	We will replace “overview the calibration results” by “overview of the calibration results”. The comment regarding the extreme event we answered in the general comments section, answer D.
comment 17	Page 9374. Line 16: Calibration equation has been presented in a previous section. Remove.
answer 17	We will rewrite the previous section. Therefore we leave the calibration equation in this section (see answer 8).
comment 18	Page 9374. Line 18: “However despite proper leveling some additional tests directly on the single buckets showed that one of the two buckets tips at a structural higher volume than the other. The cause of this difference is not clear”. Assumption of equal volume needs to be checked and volumetric calibration performed.
answer 18	We answered this in the general comments, answer B
comment 19	Page 9374. Line 25: Remove as it was mentioned at the beginning of this page.
answer 19	We will remove the sentence “Before the calibration in the lab started the tipping bucket was properly leveled.”
comment 20	Page 9375. Lines 5-9. The authors stated that the only purpose of the runoff tipping bucket gauge is to provide the mechanism for automatically emptying, and no data logger was used to record that... WHY? A data logger, in particular a real intensity one, will provide extra information needed to cross-check those values obtained from the electronic weight scales.
answer 20	We don’t see what extra information a data logger can provide. The moment of tipping is within 2 seconds accurate, so any possible cross-checking can be done by using this information. The weighing scale recordings in combination with the data treatment provide a tipping volume for every tip. This volume is checked to the expected volume of approximately 1 liter per tip.

	During the testing phase this helped us to solve a problem on one of the weighing scales. The tips volumes were approximately ¼ and ½ liter. The cause was a jammed weighing platform. During the measuring phase the funnel of the tipping bucket of the reference roof sometimes got clogged, leading to abnormal tipping volumes. We will elaborate more clearly in the text why a real intensity data logger is not required for the runoff measurements.
comment 21	Page 9376. Line 3. Use “rainfall-runoff”.
answer 21	We will change “rainfall runoff” into “rainfall-runoff”.
comment 22	Page 9376. Lines 5-10. Repetition of the weight to runoff transformation procedure.
answer 22	To avoid repetition we will remove the weight to runoff transformation procedure from section 2 to this section.
comment 23	Page 9377. Line 18. Review assumption of constant density for water.
answer 23	We will address the error we make with this assumption (see general comments, answer E).
comment 24	Page 9378. Lines 6-10. The authors refer to a storage volume in the soil. It needs to be clarified in the experimental setup description.
answer 24	We will add a short description of the green roof setup in the measurement setup description (see also specific comments, answer 11).
comment 25	Page 9379. Line 25. The reviewer does not know how to treat this result. This reviewer has serious concern with the intended used of the bathroom scale here.
answer 25	We understand the reviewers concern. We are well aware of the large inaccuracy of the bathroom scales. The bathroom scale measurements are merely to (roughly) verify the conclusions on evapotranspiration that are drawn from the rainfall-runoff measurements, and to provide a rough estimate of the stored water volume (see general responses, other aspects and budget).
comment 26	Page 9381. Lines 1- 4. The authors indicated that meteorological data (NUS weather station) and soil characteristics were used to determine evaporation. Where are the methodology and the evaporation results presented? Are they the numbers next to the arrows in Figure 10? This is confusing as the value of 6.6 mm was previously presented and explained in Figure 5 as being obtained from weight differences. Please clarify.
answer 26	The meteorological data of the NUS weather station are used to model the green roof. (general comments, answer A). This model has also to be fed with soil characteristics. The small scale weighing results (Figures 5 and 10) can be used to determine part of these soil characteristics. Next to that they can be used to calibrate the model that will be fed with the meteorological data of the NUS weather station. Some results of these measurements are presented in Figures 13 and 14 and tables 1 to 4. They are used to elaborate how these parameters can influence the weighing measurements.

	The 6.6 mm was derived from recorded weights from the small scale setup (Figure 10). For this setup a high accuracy weighing scale has been used. Since no rainfall and no runoff occurred, and the relevant parameters (radiation and temperature) at midnight in Singapore are always quite similar, the weight difference has to be caused by evaporation.
comment 27	Page 9382. Lines 1-4. "The accuracy of this approach depends mainly on the accuracy of the reading, which is in the order of few grams." This reviewer will argue that what follows in the manuscript on disturbances in runoff measurements will contradict this statement.
answer 27	We have to disagree with the reviewer on this aspect, because the relevant disturbing parameters at midnight in Singapore are constant. The graphs show that increasing radiation and temperature during day time will increase the measured weight, but they decrease in a similar amount during night time. The error induced by comparing weights at different midnights is very small.
comment 28	Pages 9382-9383. Around four pages (including Figures and Tables) were used to shown the influence of meteorological forcing on the implemented weighting system. This is an important finding as the authors documented its occurrence under field conditions. This reviewer will suggest to keep Figure 11 and 12, Tables 1 and 2, and to remove the others. Information contained in Tables 3 and 4 can be easily presented in a single paragraph. The manuscript needs this extra space to improve the Introduction and the Discussion sections.
answer 28	We agree with the reviewer that there is some "overkill" in figures and tables. We thank the reviewer for this suggestion. We will leave out tables 3 and 4. We will describe the correlation between incoming radiation, air temperature and humidity in words only. We will leave out figure 14 (humidity). Since the correlation between recorded weights and incoming radiation is highest, it seems logical to leave out figure 13 (temperature) rather than figure 12 (radiation). However 3 of the 5 scales can not be affected by incoming radiation directly and all scales can be affected by air temperature. Therefore we prefer to keep the temperature graph over the radiation graph. Next to that we can refer to this graph when we discuss the temperature – water density relation.
comment 29	Page 9383. Lines 14-16. "The error caused by these weather influences is normally very small, and that too only temporary...". This reviewer does not agree with that but this reviewer may be wrong. A 100 g difference due to radiation may result in 1.22 mm of evaporation which it can account for more than 25 % of the average evaporation for a given day (see Figure 10). The weighting system seems to be influenced by other disturbances during runoff events (changes in fluid density, climatic forcing?) and by meteorological forcing (radiation) during the dry period, thus neither is very small nor temporary. Please clarify this.
answer 29	By temporary we refer to the fact that the scales return to their original weight once the cause of the weight change is taken away. At sunrise the recorded weights increase, at sunset they decrease again, as is shown in the figures 12 – 14. The small setup the reviewer is referring to (figure 10) is not

	representative for the runoff measuring. It has been used (for a short period of time) as an evaporation experiment to determine the characteristics of the soil. When we determine the daily evaporation we compare the weights between to consecutive midnights, when radiation and temperature are almost always the same. In the runoff experiments 100 g difference coincides with 0.1 mm. Only setup 2 can come close to 100 g difference; setups 1 and 3 to 50 g; setups 4 and 5 (on other type, more expensive scales) to 20 g.
comment 30	Page 9384. Lines 1-15. The title and the first sentence are confusing since both rainfall and runoff devices use tipping bucket. Please rewrite it. At this section now, this reviewer is completely lost. Is evaporation from the tipping bucket for the runoff? In section 2.2, the authors stated that "The setup of the experiment table is such that the runoff measurement equipment can be situated bellow the opposite table. A cover around the open sites below the table helps to minimize wind and other possible influences as much as possible". According to this, they are under a cover and consequently no influenced by radiation. What drives evaporation under these conditions? Looking at real time meteorological data at Marina Bay over January 2011, this reviewer found that overnight temperatures are constant between 25-27 ^{o}C, relative humidity between 75-80%, and wind speed of around 7 km/h. This reviewer is afraid there are no ideal conditions for evaporation to occur. Please clarify.
answer 30	We refer to the flow gauge tipping bucket only. We will rewrite it to clarify this. The reviewer is right that there are no ideal conditions for evaporation. However we measure a weight loss during night and a larger weight loss during day time that can only be caused by evaporation from the tipping bucket. Sartori (1999: Critical review on equations employed for the calculation of the evaporation rate from free water surfaces) showed for a 10 hour night time evaporation values of 1.38 – 1.97 mm (kg/m^2/d). These numbers were derived for water temp. 25 ^{o}C, air temp. 20 ^{o}C, wind velocity 2 m/s and relative humidity 100 %. Of course these circumstances are not exactly similar to our situation, but it clearly indicates possible night time evaporation.
comment 31	Page 9384. Lines 7-8. The authors refer to "the runoff of the concrete roof". What is it?
answer 31	That is our reference roof setup, where only a concrete layer is applied. We will change "the runoff of the concrete roof" into "the runoff from the reference roof setup"
comment 32	Page 9384. Line 23. Change "7.62 mm" by "76.2 mm".
answer 32	We will change "7.62 mm" into "76.2 mm".
comment 33	Page 9385. Lines 8-11. A 3% error for one tip per minute (or 12 mm/h intensity) is an unreliable figure. What causes that error? With this intensity, individual drops of water will be filling the bucket, reducing the half tip time error to a minimum. Could it be also the result of the release on water using the syringe for such a low intensity? That small difference may indicate the need for a proper adjustment of the

	calibration stop screws for each bucket.
answer 33	The reviewer is right. A proper factory calibration should result in zero error when individual drops are filling the bucket. In stead of trying to adjust the calibration stop screws we just used the value of 0.2068 mm for a single tip per min that was a result of our calibration. Larger rainfall intensities require larger tip volumes anyhow (as was shown in our calibration). Literature review on green roof monitoring showed us that often the factory values (for a single tip) are taken for granted, also for increasing rainfall intensities.
comment 34	Page 9385. Line 25. Check citation format for Devine and VanWoert et al.
answer 34	The correct citation format of these two references is applied earlier in the same paragraph. We assumed that once in every paragraph is enough. Obviously we assumed wrong. We will change (twice) Devine into Devine (2009) and VanWoert et al. into VanWoert et al. (2005).
comment 35	Page 9386. Lines 1-9. It is not clear what the authors try to say here. A real intensity logger stamps the date and time corresponding to each tipping. In terms of memory, it may use more of it. For example for a 0.2 mm bucket, 250 records (or lines) will be used for a 50 mm rainfall event over 1 hour, but only 60 records if the logger is set for 1 min interval. All existing data loggers have clock issues over long period of time, but the error will be negligible over consecutive tips over short period of time, thus the “real intensity” can be measured. Again, frequent resetting of the logger timer decreases the clock issue. Since this is the same issue for the clock of the weight data logging system used for the tipping runoff gauge, this reviewer does not see the point for those comments placed in the paragraph.
answer 35	We try to say that on the long run (with lots of dry minutes) a real intensity logger is better regarding the memory use, since it only records actual tips. During a severe rainstorm (like in the example of the reviewer) this is temporarily opposite. We agree with the reviewer that during severe rainstorms such a logger will provide additional information on the rain intensity. We simply warn for bogus accuracy during (very) light rainfall intensities. In fact during our setup phase we looked for tipping bucket rain gauges with a data logger that records date and time of single tips. We could not find this combination available (within our limited budget). Regarding the tipping of the runoff gauge we like to refer to specific comments, answer 20.
comment 36	Page 9386. Lines 10-19. Discussion about wind and other effects on rainfall measurements which have no relevance to the lab calibration performed in this work. There are out there practical solutions for each of the listed problems in order to minimize their influence on the rainfall measurements. The authors may consider a look into the work by Grzegorz J. Ciach (2002).
answer 36	We like to thank the reviewer for this suggestion. Regarding the wind influences, we recently (November 2010) changed the setup of the rain gauge. We removed it from its original position and placed it in the centre of the reference roof setup. That way the wind

	effect is minimized. The rain gauge disturbs the runoff from the reference roof in a minimal way. The tipping bucket water is discharged on the roof itself. The delay is very small. 3% of the area discharges later (a few seconds during intense rainfall). The new setup is more in line with the rain gauge setup used by Van den Eertwegh (2002). Since we changed the rain gauge setup we didn't record more runoff from the reference roof than rainfall. So it seems to improve the total setup.
comment 37	Page 9386. Lines 20-25. First sentence of this paragraph is confusing. What is the "reference roof runoff measurement"? Since this is ongoing study, this reviewer will recommend the authors to wire all the tipping gauges (rainfall and runoff) to real intensity data loggers.
answer 37	We mean to say that green roof runoff should be compared to runoff from a reference roof rather than to rainfall. By "reference roof runoff measurement" we mean the recorded runoff from the reference roof setup. We will change this sentence into: That implies that comparing measured runoff from a green roof to measured runoff from a reference roof will lead to much more accurate results than comparing measured runoff from a green roof to measured rainfall. We like to thank the reviewer for the recommendation. We will consider it regarding the rain gauge. Regarding the flow gauges we like to mention that they are already connected to an (almost) real intensity data logger (see specific comments, answer 20).
comment 38	Page 9387. The first paragraph is just a compilation of previous studies using different gauging setups and it should be moved to Introduction to put into context the present work. Unfortunately, the authors seem to confuse "capabilities of a device" with "scanning time and recording time" of a given method. Pressure sensor readings have the advantage of being compensated by temperature and atmospheric pressure (avoiding some of the issues experienced by the weighting system) and scanning and recording time can be set at 3 sec intervals which will provide enough data (20 points per minute) for high temporal resolution. The authors certainly have explored a novel methodology that may provide 1 sec time interval information but the method needs further testing of the assumption and external influences.
answer 38	We will rewrite the paper in a more traditional format (see general responses, paper format) and therefore move the literature review to the introduction. We will separate resolution from device capabilities more clearly. The device we are referring to allowed determination of 0.5 mm of water runoff from the roofs. We are not questioning the device that was used, but we are aiming for a smaller allowed determination than 0.5 mm.
comment 39	Page 9388. This reviewer does not have any observation regarding the water balance approach followed by this and previous works to account for evapotranspiration. However, the error introduced by radiation is of the same order of magnitude than the experienced weight differences used for evaporation estimates. Finally, previous studies should be used in the Introduction to provide the right context for the present work.

answer 39	Regarding the small scale (high resolution) measurements the reviewer is correct that weight errors caused by radiation are relatively large when the continuous measurements are considered. For comparing different "midnight" weights we have to disagree (see specific comments, answers 26, 27 and 29). Regarding the bathroom scale measurements we are well aware that the errors can be large, way larger than weighing lysimeters. However we used them to get an impression of the evapotranspiration and storage during periods between the ends of two green roof runoff events. Of course a weighing lysimeters are much more accurate for this.
comment 40	Page 9389. Conclusions are very poor and reflect the fact of lack of clarity on what this work is about. For example, the authors again concluded about the effect of wind on rainfall measurements subject for which they have not done anything in the present work (Lines 12-14).
answer 40	We will focus the conclusions more on what this work is about (new way of measuring runoff, compare green roof runoff to reference roof in stead of to rain), and leave out conclusions regarding the wind effect on rain measurements.
comment 41	Table 1 and Table 2 captions: Misspelling " Februari 2010"
answer 41	This will be changed in "February 2010"
comment 42	This reviewer suggests to remove Figures 13 and 14 as well as Tables 3 and 4.
answer 42	We will remove Figure 14 and Tables 3 and 4. Regarding Figure 13 we prefer to remove Figure 12 on radiation. The reason for this we tried to explain in specific comments, answer 28.

Preliminary versions

Abstract
This article describes the measurement setup that is used for green roof experiments in a tropical environment, the required data treatment to obtain reliable values of rainfall – runoff and how to deal with external disturbances that can influence the experiment results. High resolution rainfall – runoff measurements to identify, understand and properly model the relevant runoff processes in a green roof require both tailored equipment and data treatment. A runoff measuring setup is developed that can accurately quantify the runoff up to 6 liter per minute, and has a high resolution in both time and volume. It is concluded that rainfall – runoff from a green roof should rather be compared to rainfall – runoff from a reference roof and not to rainfall, like it is normally done.

Keywords
High resolution, Green roof experiments, Rainfall – runoff, Tropical conditions.

Introduction
Various methods have been used to measure rainfall – runoff from green roofs. Moran et al. (2004) collected the runoff data with a V-notch weir box with a level sensor. Data were recorded in 5-minutes intervals. Mentens (2003) and VanWoert et al. (2005) used rain gauge tipping buckets with a 5-minute interval for runoff. Stovin et al. (2007) recorded the runoff data by means of a collection tank with a high resolution pressure transducer. Though stated to use 5-second intervals, data in the paper were presented at 5-minute intervals. Uhl and Schiedt (2008) used collection tanks with a swimmer system for water level measurements to measure the runoff. EPA (2009) used 55 gallon plastic barrels with a pressure transducer which allowed continuous measurement of the water level in the barrel. Pressure data was recorded every 5 minutes. The pressure transducer systems have a sensitivity which allows determination of approximately 0.02 in. (0.5 mm) of water runoff from the roof area. They stated that the sensitivity was more than adequate for assessing the total volume of the storms, but did introduce some uncertainty and variation in time series analysis or instantaneous results, as the transducer might toggle between values or experience drift.

When green roofs are applied, they replace conventional roofs. To quantify the effect of a green roof it therefore seems more logical to compare the runoff from a green roof to the runoff from a conventional roof. In addition rainfall measurements often lead to underestimation of the rainfall depth, either due to wind effects (Sevruk, 1996, Hsu and Guo, 2005), or due to increasing tipping bucket losses at increasing rainfall intensities (Marsalek, 1981, Devine, 2009). Remarkably, all studies mentioned above determine green roof runoff reduction by comparing the green roof runoff to rainfall that was obtained from tipping bucket rain gauges with a recording interval of 5 minutes. Only VanWoert et al. (2005) mention that the accuracy of the rain gauge decreases with increasing intensity, but like the others they don't mention any data adjustment for higher rainfall intensities, nor for wind effects.

These research projects all focused on long term effectiveness of green roofs and the measuring equipment that was used generally fits this purpose. The rainfall – runoff measurement setup described in this paper is used to measure runoff dynamics with a high resolution both in time and in volume in a tropical environment. The latter implies a large variation in rainfall – runoff intensity. The way of monitoring and the size and shape of our experiment tables have been adjusted to each other. In addition practical aspects such as the load capacity, roof access and the costs per setup were taken into account.

Rain gauge calibration

Before installing the rain gauge we calibrated it in our laboratory, because tipping bucket rain gauges always have an accuracy that is decreasing with increasing tip frequency (Marsalek, 1981). Main reason for this decreasing accuracy is that it takes some time for the bucket to tip (Devine, 2009). During half of the tipping time rainwater is flowing into the filled bucket, but actually belongs to the next tip. The faster the flow rate the larger this error is. Calibration is also required because tipping buckets generally have a small deviation from the tipping value provided by the manufacturer. This deviation can be adjusted by using the two calibration stop screws (one for each bucket). We used a different approach, because rainfall in Singapore is often very intense.

The rain gauge has been calibrated in a lab, by performing several tests. The rain gauge was properly leveled before the start of the calibration. The first test was a volumetric verification, carried out by dripping water from a syringe directly into the bucket, until it tipped. The volume difference in the syringe was measured with an accuracy of 0.1 ml. This test was carried out three times for each bucket. This resulted in average tipping volumes for the two buckets of 6.50 ml and 7.23 ml, i.e. 0.2004 mm and 0.2230 mm for an 8 inch rain gauge. The difference between the two tipping volumes is a little over 10 %. Not using the calibration stop screws implies an error of a little over 5 % for a single tip. However, the higher the number of tips the smaller this error gets. In the long run the error caused by this difference becomes negligible.

The second test was carried out by emptying exactly 100 ml of water through a syringe into the rain gauge funnel at different speeds. During 20 tests every tip was recorded using a stop watch. For calibration only the total number of tips, the time of the last tip and the rest volume in the tipping bucket are used. The pouring is done manually, so the pouring speed is not exactly constant during the entire pouring time, as will be when real rainfall is recorded. The average tipping frequency of the 20 tests varies from 2.01 to 27.69 tips per minute. To determine the capacity of the rain gauge an extreme test was carried out. 100 ml of water was emptied from a measuring jar within a few seconds into the funnel of the rain gauge; resulting in an average tipping frequency of 67.50 tips per minute. Based on these results a linear relation between the average tip frequency and the average tip volume has been determined (Figure 5).

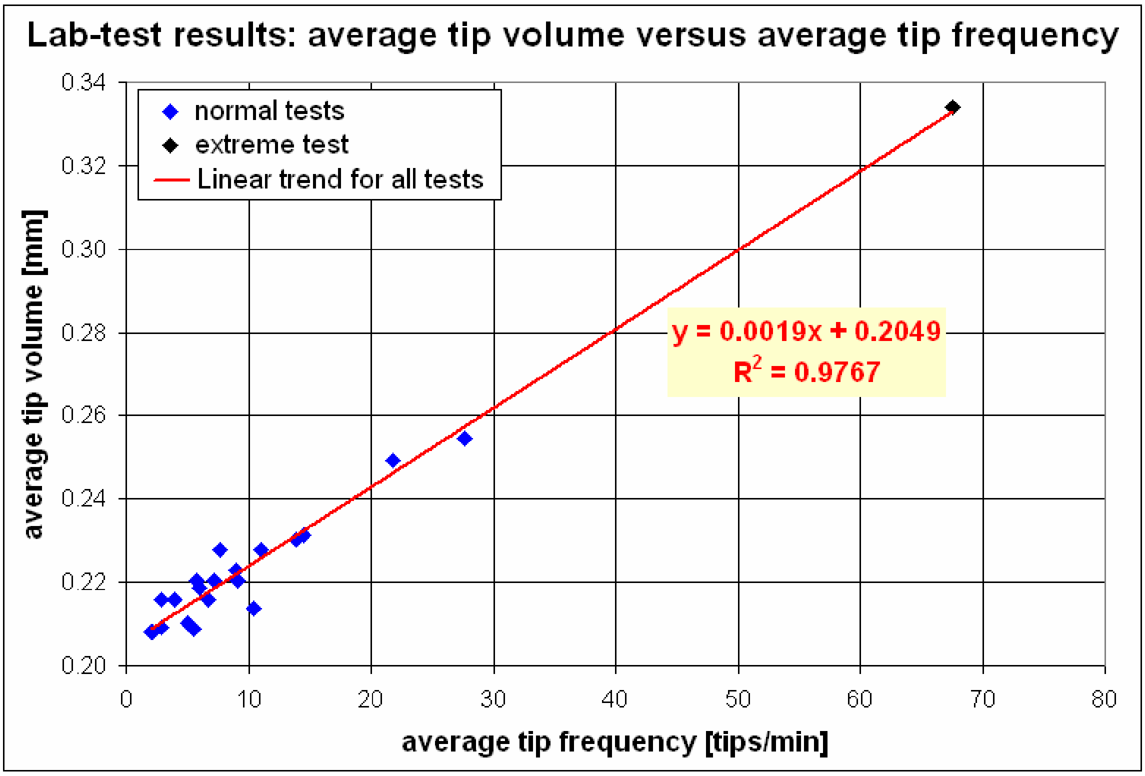

Figure 5 Result of the rain gauge calibration

As a result of this the rain gauge recordings will be adjusted according the following conversion rule:

Rainfall [mm/min] = NRT * (0.2049 + 0.0019 * NRT)

In which NRT is the number of recorded tips in a single minute

A minute with a recorded number of 18 tips (the highest recorded rate so far) implies a rainfall of 18 * (0.2049 + 0.0019 * 18) = 4.3038 mm during this minute. That is almost 16 % more than 18 times a single tip volume of 0.2068 mm.

A third test on the tipping bucket rain gauge was carried out at a much later stage, long after the green roof experiments started, when during several events runoff from the reference roof exceeded the rainfall. During this test the tipping registration mechanism of the rain gauge was checked by manually tipping the bucket over and over again and counting the number of tips. The tip-frequency during these tests varied between 6 tips per minute to more than 100 tips per minute. The tests were conducted with a dry tipping bucket magnet, with a wet tipping bucket magnet, and even with a piece of paper that covered the magnet. Several thousand tips were all recorded by the data logger. As a result of this we concluded that other effects like wind, wetting, evaporation and splashing effects must be the cause of this obvious underestimation of the actual rainfall.