14 November 2014

Dear Editor Prof. Andreas Stohl,

We hereby submit the revised manuscript entitled "Detailed source term estimation of the atmospheric release for the Fukushima Daiichi Nuclear Power Station accident by coupling simulations of atmospheric dispersion model with improved deposition scheme and oceanic dispersion model" by Genki Katata et al.

First of all, we apologized that we took very long time for revising the paper to answer to all reviewer comments and suggestions. We did not upload the marked-up manuscript because we thought that tracking our revisions (for more than 100 comments) would just confuse the editor. Instead, we chose to upload the precise point-by-point response files so that the editor can clearly identify what we changed in the revised version.

I sincerely appreciate your consideration our manuscript for publication.

Best regards,
Genki Katata, on the behalf of all authors

**Title: Detailed source term estimation of the atmospheric release for the Fukushima Daiichi Nuclear Power Station accident by coupling simulations of atmospheric dispersion model with improved deposition scheme and oceanic dispersion model**
**Authors: G. Katata et al.**
**MS No.: acp-2014-330**

## Author response to reviewer comments

## Response to Ref.1

The authors present an ambitious article on the analysis of the Fukushima accident in the light of all the work that has been done on the subject. The paper addresses various topics: source term estimation and analysis of the release events; verification of the source term at local, regional and global scales. They partly raise the issue of the sensitivity to the atmospheric dispersion model, to the meteorological data and to the deposition parameterizations. The analysis of the accident (impact of the release events, connection between the release events and the events that occurred in the plant, comparisons with observations) is highly comprehensive. The supplement of the paper is also very useful.

➔ We appreciate your careful reading and giving positive comments on our paper. We tried to respond to suggestions in detail as much as possible, but we sometimes could not bring the whole sentences from the manuscript to this letter because they are too long to cite in the letter. We showed the chapter, section, or subsection number, so please also see the revised manuscript.

The uncertainties associated with the source term estimations due to the assessment method and to the meteorological data need sometimes to be refined and reminded in the text (section 3.1 in particular). The part concerning the importance of improving the modeling of the deposit is less convincing and need to be further developed (without increase the paper size) or addressed more fully in another paper. Given the current state of the study, the authors have to be much less clear-cut on the advantage to use a more sophisticated deposition scheme.

➔ As replied to all of your comments below, we revised the manuscript to solve above concerns of yours within a shorter paper size. We hope the advantage of the new source term and deposition scheme is much clear in the revised version.

The paper is very long. The authors can probably shortening some parts. Some figures may be moved to the supplement of the paper. Some suggestions are given.

➔ On the basis of all reviewer comments, we made the following revisions to reduce the paper length summarized as:

- In Introduction, the sentence after 'First, …' of ln.5/ P.14730 was shorten.
- Chapters 3 and 4 are reconstructed. 3.1.1-3.1.9 which describes the source term and the motion of plume during March was revised and only the periods which were different from the previous paper (Terada et al.

2012) were remained. Related this revision, Figs. 5 and 6 were deleted. As for chapter 4 (Discussion), section 3.2 (Verification of source term) was moved to section 4.1. Section 4.1 (Comparison of source terms) was moved to section 4.2. Finally, section 4.2 (Role in deposition process) was moved to Supplement.

- The section for “Validation using several models” was shortened. Related this revision, (old) Figs. 18, 20, and 22 and Tables 7, 8, and 10 were extracted.
- Conclusion was also shortened based on reviewer’s comments.

At the same time, we needed to include more sentences to reflect all reviewers’ comments to the manuscript (e.g., Meanwhile, Fig. S3 was moved to the main text as (new) Fig. 12). However, the page numbers and numbers of figures and tables were still reduced compared with our previous paper.

<u>MAJOR COMMENTS</u>

<u>Section 2.2: Reverse estimation method over the land</u>

P14735 What do the authors mean when they explain that they used peaks values from continuous time series of air concentration? Why is it better to do so than to use the full set of data?

➔ The comparison of peak values is carried out for JAEA-Tokai data, because time series data are only available at this place. The air concentration data from JAEA-Tokai is one hour averaged one. When we use the reverse estimation method, “Since the uncertainty of model simulation is the primary cause of the discrepancy in the spatiotemporal distribution of plume between the measurements and simulation results, the above procedures cannot be applied systematically, and the correction of this discrepancy by ‘expert judgment’ is necessary to reduce the impact of model uncertainty on the source estimation. The process is to check all available measurements to see if the plume is reproduced appropriately or not for comparison with the measurements, and to determine if the discrepancy is caused mainly by errors in the calculated wind direction. If the plume flow direction is clearly different from the measured wind direction, the calculated plume is rotated to match the measured wind direction and Eq. (1) is applied. The use of peak values corrects any discrepancy in the timing of the arrival of the peak air concentrations between the measurement (JAEA-Tokai) and simulation. We assume that the peak values of the measurement and simulation are comparable even though the timing or temporal pattern of the arrival of the peak is different because the central plume axis passes across the sampling point differently between the measurement and simulation” (new subsection 2.1.1).

P14735 With dose rate measurements, only the observed air dose rate from ground shine is used to assess the source term. If so, how the timing of releases (beginning and end) is estimated?

➔ In this study, we divided the release interval and looked for when and where the specific segment of plume increased air dose rate, and then found the appropriate observation data which can be used for the source term estimation for the certain plume. Thus, the timing of releases should be clear. The ground-shine was used for the source estimation on 12 and 15-16 March. For the case of 12 March, the release period for wet venting of Unit 1 was determined from the decrease of pressure of drywell and the release by hydrogen explosion is assumed almost instantaneous. For the case of 15-16 March, the release period was basically divided with every hour and the source term for each plume was estimated by the method as mentioned above. The descriptions how to find the appropriate observation and how to determine the release period of the segment of plume are available in (new) subsection 2.1.1 and 2.3.3.

The isotopic composition is assumed to assess the source term by using observed dose rate from ground shine. Release rate for noble gazes is not assessed and computed dose rate signal does not take into account the contribution of noble gazes to the plume dose rate. Nevertheless Chiba observations showed that for some release events a large part of the total dose rate was due to the noble gazes contribution when the plume was detected. How do the authors interpret the comparisons between total air dose rate (including plume contribution) simulated and observed shown Figures 16 and in the supplements? What is the impact of their assumption on the source term assessment and its evaluation?

➔ In the aspect of source term estimation, we did not consider the effect of noble gases because we used ground-shine for our source estimation from air dose rates. For validation, (new) Fig.11a showed the comparison of air dose rates of 18 March when the plume flowed toward the Pacific Ocean. Hence, the air dose rates over the land should originate from only the ground-shine of deposited materials except for noble gases. The comparison for validation in (new) Figs. 6a and b, and Fig. 12 was done at a slow decrease in the dose rates after the peak (ground-shine). These points are added to sections 3.1 and 3.4, and subsection 4.1.1.

<u>Section 2.3: inverse estimation method over the ocean</u>

Except if I missed something, the method used to assess the source term over the ocean is not based on inverse modelling techniques. To avoid any confusion, the authors should not call their method “inverse estimation method”.

➔ According to the reviewer’s comment, we changed the description for the method used to assess the source term over the ocean from “inverse” to “reverse”. Also, at the first part of (new) chapter 2, we defined the reverse and inverse methods with merit and demerit, and also described why we choose the reverse method for the FNPS1 accident case as “A reverse method evaluates the release rates of radionuclides by comparing measurements of air concentration of a radionuclide or dose rate in the environment with calculated one by atmospheric transport, dispersion and deposition models (ATDM) for a unit release of a radionuclide. The release rate is estimated by the ratio of the measurement to calculation result. The merit of the reverse method is that the comparison can be made with one or more independent data points. For example, the minimum number of data points needed is only one and the measured data used for the estimation can change with time from air concentration to air dose rate and vice versa. The demerit is that this simple comparison without consideration of the uncertainty of the ATDM results may cause the large errors, and, consequently, expert judgment is essential to correct the discrepancy between the measurement and calculation.”

This part of the paper has to be improved: the goal and the method have to be clarified.

➔ The goal and method were clarified in (new) subsection 2.1.2. Also, the suffix and equations are revised according to the comments from other reviewers.

Section 3.1 Source term estimation and local-scale dispersion analysis

The reliability and uncertainties of the meteorological data should be given for the various release events.

➔ Although we realize the importance of the evaluation of the reliability and uncertainties of meteorological data for various events, the detailed analysis on this matter needs more space in paper and should be considered in future work. In the revised paper, we partially discussed this issue in subsection 4.1.1, Conclusions, and Supplement.

What monitoring data are used to reconstruct each release events? Their number has to be given (it could be given Table 5). The relevance of the various emissions has to be discussed. This requirement is at least needed for the release events showing the main discrepancies between the actual study and the previous one.

➔ We revised (new) Table 6, in which the monitoring data used for estimation is shown as monitoring locations in (new) Figs. 2 and 3. Also, in sections 3.1-3.5, we described which data are effective to determine the new source term.

What are the specific reasons for the new release assessment (especially on March 15-16)? This point is partially discussed Section 4.1 and need to be completed.

➔ "The events in the reactors (TEPCO, 2011a, 2012; Tanabe, 2012) are also shown in Fig. 5, but it is not clear from the reverse estimation that the events written in Fig. 5 mainly caused the atmospheric releases, particularly after 15 March" (the first paragraph of new chapter 3). Meanwhile, it is clear that the new release assessment on 15-16 March is achieved by both new observation data, e.g., the monitoring post data from Fukushima Prefecture and the improvement of deposition scheme (new section 3.4 and subsection 4.1.1). By this finding, the overestimation of deposition in the south area of Miyagi Prefecture in the previous work was clearly improved when using the new source term (new subsection 4.1.1).

Section 3.2

The main release events that are different from the previous study could be analyzed in more details. Does it give a better agreement by comparing simulations with dose rate measurements and/or air concentrations measurements? Monitoring dose rate comparisons shown in the supplements may be used to explain the impact of the new source term compared to the previous one.

➔ Considering the reviewer's comment, we concentrated on the analysis of release events that are different from the previous study in chapter 3. In particular, for the releases from 15-16 March, we show the comparison in air dose rate at selected monitoring sites affected by large deposition events (Hirono, Kawauchi, Fukushima, Iitate, Kawafusa (NW of FNPS1), and Yamada) in the maintext (new Fig. 12). We hope now the improvement due to new source term is much clear.

Regional deposition

Authors claim that "both improvements resulted from the enhancement of the scavenging coefficient by including in-cloud scavenging in the modified wet deposition scheme" the demonstration is not conclusive and it is difficult to precisely identify what is the specific contribution of the new deposition scheme. The authors should compare simulations done with the new source term and the previous deposition scheme with simulations done with the new configuration and source term. Those simulations could be compared in the various tables.

➔ Considering the comments from all reviewers, we carried out four cases of simulation in terms of deposition distribution of Cs-137 using the combinations of (1) old source term (Terada et al., 2012) and original WSPEEDI-II, (2) new source term and original WSPEEDI-II, (3) old source term and modified WSPEEDI-II and (4) new source term and modified WSPEEDI-II. By comparing the results from (1) and (3), the effect of model improvement on deposition distribution was examined and then, by comparing (3) with (4), the effect of source term improvement to the distribution was investigated. The results were described in subsection 4.1.1.

Moreover, it seems surprising that the in cloud scavenging has a large impact. Indeed, the plume was probably situated in the lower layers of the atmosphere at the regional scale and below cloud scavenging may have been dominant.

➔ We did not consider below-cloud scavenging in the previous manuscript because it was found to have much less significant impact as mentioned in many prior studies (A6 in (old) Appendix). Even though it was less significant, however, we still needed to show that the contribution of below-cloud scavenging to the total wet deposition. In the current simulation, we computed the below-cloud scavenging process and estimated its contribution to the total wet deposition as the difference between the modeled depositions with and without below-cloud scavenging divided by that with below-cloud scavenging (%): $(D_{with\_bl} - D_{wo\_bl}) / D_{with\_bl} \times 100$. The wet deposition amount due to the below-cloud scavenging was highest in the northwest vicinities of the power plant, 10 kBq m$^{-2}$. We investigated the reason why the contribution of below cloud scavenging was such low in the northwest region by drawing horizontal-vertical cross section of the plume in the night of March 15 when the highest contamination occurred there. In the event, the cloud base height (0.01 < LWC < 1 g m$^{-3}$) was very low near to the ground so that most of radionuclides scavenged by in-cloud processing and the contribution of below-cloud scavenging was relatively low. Amongst the whole regional-scale model domain as well as local-scale domain, the highest contribution of the below-cloud scavenging was smaller than 1%. The model formulation of the below-cloud scavenging process and the quantitative comparisons between the simulated in-cloud and the below-cloud scavenging coefficients were given in (new) Appendix. Further details of the analysis are beyond the scope of this study, which is the estimation of the emission inventory, and so those will be discussed in the future works.

Comparisons done table 6 are not homogeneous. It is not always the same simulations that are compared with the “New-land” one. It has to be more homogeneous.

➔ We completely revised the table for four WSPEEDI calculation cases (combinations of the original/modified WSPEEDI and Terada/new source terms) (new Table 7) so that readers can understand the impact of new source term.

<u>Local air dose rate</u>

A table giving the statistical indicators for air dose rate comparisons should be added.

➔ (Old) Table 6 included the statistics compared with local-scale air dose rates. Furthermore, in (new) Table 6, statistical indicators were provided for four simulation cases (combinations of the original/modified WSPEEDI and Terada/new source terms) so that readers can readily understand the main result in this study.

Section 4.1

Chapter 4.1 could be moved to the beginning of Section 3 in order to better highlight the specificity of the new source term estimation.

➔ (Old) Chapter 4 (Discussion) was completely reorganized to focus on new source term estimation.

The authors claim that the release of March 15-16 is assessed because of the new data set: what data are useful to reconstruct this event?

➔ The advantage of new dataset was described in (new) chapter 3.

I do not believe that the modified wet scavenging scheme could explain the new timing of the release event. It can help to decrease the release rate on March 15 pm but it cannot explain the increase of the release rate in the evening.

➔ The sentence was inappropriate. We corrected the part as the new timing of release events was found out by using new monitoring data in section 4.2 as "These were particularly effective to find this release and determine the timing and release rates."

Section 4.2

As previously said, it is difficult to precisely identify what is the specific contribution of the new deposition scheme and I do not believe that the authors should end their paper with this section. This discussion could be dispatched partly in section 3.1 and partly in section 3.2.

➔ This section was extracted from the maintext (moved to Supplement) so that we focused on the impact of new source term (subsection 4.1.1).

What is the relative contribution of below cloud and in-cloud scavenging (especially at the local scale)?

➔ As replied above, the wet deposition amount due to the below-cloud scavenging was highest in the northwest vicinities of the power plant, 10 kBq $m^{-2}$. We investigated the reason why the contribution of below cloud scavenging was such low in the northwest region by drawing horizontal-vertical cross section of the plume in the night of March 15 when the highest contamination occurred there. In the event, the cloud base height ($0.01 < LWC < 1$ g $m^{-3}$) was very low near to the ground so that most of radionuclides scavenged by in-cloud processing and the contribution of below-cloud scavenging was relatively low. Amongst the whole regional-scale model domain as well as local-scale domain, the highest contribution of the below-cloud scavenging was smaller than 1%.

The authors should be less conclusive on the beneficial contribution of the new deposition scheme considering the various uncertainties (meteorological data, iodine speciation…). Moreover Table 7 shows that the model to data comparison may be less sensitive to the MLDPO deposition scheme than to the meteorological conditions (NAME simulations) and to the source term.

➔ We agreed with the reviewer comment, and revised the description of (new) subsection 4.1.1 so that the beneficial contribution of the new deposition scheme is less conclusive considering the several uncertainties.

The authors present the fog deposition scheme as an important improvement. What about the quality of the fog and drizzle simulations with MM5? What about the fog observations? Light rains are not detected with radar observations. Are they with rain gauge?

➔ According to the comments from other reviewers, we excluded (old) section 4.2 about deposition processes (and moved it to Supplement). However, you can find the simulation result of cloud liquid water content (LWC) at the ground station (Fig. S8). MM5 clearly underestimated LWC event which caused the underestimation of fogwater deposition. Yes, the data referred in the previous manuscript was radar data merged with rain gauge data. More detailed analysis should be provided to the future study.

The relevance of the precipitation data should have been discussed before the end of the paper since it has a huge impact on the release assessment. For instance, what is the impact of the over-estimation of the rain data on March 20?

➔ We partially addressed the relevance of the precipitation data and the release assessment in subsection 4.1.1 and Conclusion. Further analysis is beyond the scope, but some discussion of deposition process is available in (new) Supplement.

<u>Appendix: below cloud scavenging</u>

Nucleation scavenging rate is a process to be considered for in cloud scavenging and not below cloud scavenging. What is the point of the authors?

➔ Yes, we reworded in cloud scavenging by nucleation scavenging in the previous manuscript. In 3-D chemical transport models, in cloud scavenging (nucleation scavenging) is dominant to below cloud scavenging for the removal of hygroscopic submicron aerosols. That was what we meant in the previous manuscript.

How the below cloud scavenging is parameterized since you do not consider aerosol-hydrometeor coagulations scavenging? What is the relative contribution of below cloud and in-cloud scavenging at the regional scale?

➔ In the previous manuscript, we did not consider the below-cloud scavenging because it was found much less significant in our simulation. As replied above, the contribution of below cloud scavenging was smaller than 1% using the revised WSPEEDI-II.

At the local scale, the plume may be situated below the cloud. Therefor below-cloud scavenging cannot be neglect compare to in cloud scavenging.

➔ As replied above, we investigated the reason why the contribution of below cloud scavenging was such low in the northwest region by drawing horizontal-vertical cross section of the plume in the night of March 15 when the highest contamination occurred there. In the event, the cloud base height ($0.01 < LWC < 1$ g $m^{-3}$) was very low near to the ground so that most of radionuclides scavenged by in-cloud processing and the contribution of below-cloud scavenging was relatively low.

Paper organization

Section 2.4.1: the reverse estimation method is partially described in section 2.4.1 instead of section 2.2. The observations used in the study were partially described in section 2.2.

See suggestions for Section 4.

➔ Considering all reviewer comments, we completely re-organized the contents of all sections.

MINOR COMMENTS

Section 1: Introduction

P14730 the argument developed following « First, the estimation… » has to be clarified. Too many things are discussed.

➔ This part was revised in (new) Introduction.

When explaining the source of discrepancy they need to add the uncertainties in the meteorological data (wind, rain…).

➔ Although we realize the importance of the evaluation of the reliability and uncertainties of meteorological data for various events, the detailed analysis on this matter needs more space in paper and should be considered in future work. In the revised paper, we partially discussed this issue in subsection 4.1.1, Conclusions, and Supplement.

Section 2.2: Reverse estimation method over the land

Assumptions regarding the ratios of I2, CH3I and particulate iodine have to be specified together with the isotopic composition of the release. The authors need to evoke the strong uncertainties due to the isotopic composition of the release and a fortiori of speciation of the iodine, the behavior of iodine into the atmosphere. This discussion can be done section 2.4.4 if more appropriate.

➔ The ratio of gaseous and particulate iodine is determined from the measurement at JAEA-Tokai and the ratio of $I_2$ and $CH_3$ is assumed based on reference, RASCAL4.0, because no observed data on the ratio of $I_2$ and $CH_3I$. The determination of the ratio of $I_2$, $CH_3I$ and particulate iodine mentioned above has strong uncertainty. Because the deposition manner of these three types of iodine is different in environment, the estimation of source term for iodine is affected by this uncertainty. This was written in subsection 2.3.3.

Section 2.4.4: Simulation settings

What meteorological data were used and when? The authors should precise the method they used to choose the more appropriate meteorological data for each release event.

Are the meteorological data different from the previous study? What are the differences?

➔ As description was obscure, we revised as "Two sets of meteorological input data, a Grid Point Value (GPV) of the Global Spectral Model for Japan region (GSM) and the Meso-Scale Model (MSM) provided by the Japan Meteorological Agency (JMA) are used for initial and boundary conditions of MM5. MSM which covers over Japan with finer resolution is adopted to the reverse estimation over the land and GSM over the globe to the estimation over the ocean" in subsection 2.3.2.

<u>Section 3.1</u>

Section 3.1 is very interesting but it is sometimes difficult to discriminate between what is known for sure and what is due to the analysis of the results/model outputs. For example, P14745 "the light rain or drizzle". Is it observed? The text has to be carefully re-read in order to avoid any ambiguity.

➔ We revised words from "the light rain or drizzle bands" to "rain bands" observed by rain gauge at Koriyama (Fig. S1). We also revised the whole text to avoid ambiguity of wording.

<u>Section 3.2</u>

Statistical indicators must always be the same in the various tables and in the text. For instance, p 14752 FA 10 is used; p 14754 FA 5 is used. You should use always FA5 for instance.

➔ To make the paper shortened, we deleted Tables 7-8 for WMO model calculations in the revised version. In addition, Table 6 has been fully revised with the statistics of CC, NMSE, FB, and FA2/FA5/FA10 to demonstrate the difference between four WSPEEDI simulation cases (combinations of the original/modified WSPEEDI and Terada/new source terms) in subsection 4.1.1.

Validation using several models: this part can be shortened

➔ The text is shortened to exclude the discussion using NAME model, which is now present in supplement.

<u>Section 5</u>

Modifications have to be done in accordance with the previous remarks (uncertainties, impact of the new cloud scheme…).

➔ Conclusion was modified to show the potential uncertainties with the new source term.

<u>Tables</u>

A table similar to table 3 could be added for dose rate observations used in the reverse method.

➔The new table (Table 3) for air dose rate was added to the revised version

Table 5: description of the last column is missing. Does it give what monitoring data are used to assess the source term?

➔Thank you for your suggestion. We revised (new) Table 6.

Tables 6-7-8: please give the same list of statistical indicators (add FA2, FA5, FA10 in tables7-8 and NMSE, FB, FMS, KSP, Rank in table 6).

➔ As replied above, we deleted Tables 7-8 for WMO model calculations in the revised version to make the paper shortened. In addition, Table 6 has been fully revised with the statistics of CC, NMSE, FB, and FA2/FA5/FA10 to demonstrate the difference between four WSPEEDI simulation cases (combinations of the original/modified WSPEEDI and Terada/new source terms) in subsection 4.1.1.

<u>Figures</u>

Generally Figures are too small.

➔ All figures were modified so that the readers could readily understand them.

Some figures are not essential and may be removed if the paper is too long for publication. For instance
Figure 1 can be suppressed.
I am not sure that Figures 5-6 and 23 are required.
Figures 8 are too small to be useful.
There is a problem with the blue curve on Figure 11c.

➔ According to your comment, many figures (e.g., old Figs. 5, 6, and 23) were extracted from the main text. As mentioned above, (old) Figs. 8 and 11 were also modified.

Bands within a factor of 10 have to be added on Figures 18-20-21-24.

➔ According to your comment, bands within a factor of 10 are added to (new) Figs. 19 and 21.

<u>Appendix</u>

The authors should give more information on the initiation of the various parameterizations and the rain threshold used.

➔ We tried our best to describe each variable used in the equations of (new) Appendix in the revised manuscript. The rain thresholds were added in Appendix. The in-cloud scavenging is activated in a model grid where cloud water mixing ratio is higher than $10^{-6}$ (kg $kg^{-1}$) and the surface precipitation intensity is larger than zero. The below-cloud scavenging is activated in a model grid where each settling hydrometeor mixing ratio is higher than $10^{-9}$ (kg $kg^{-1}$) and the surface precipitation intensity is larger than zero.

The authors should talk about Iodine particulate instead of restricting it to particulate I-131.

➔We revised the word “particulate I-131” to the general form “particulate iodine” in Appendix section.

L22 p14770 has to be modified.

➔ We revised the word “particulateI” to “particulate iodine”.

## Response to Ref.2

The manuscript presents the new source term estimation of I-131 and Cs-137 released into the atmosphere from the Fukushima Dai-ichi Nuclear Power Station (FNPS1) in Japan by inversion analysis combining measurement data and offline coupling model of the atmospheric and oceanic dispersion models. Also, the manuscript evaluates the new source term by comparing the simulation using different atmospheric dispersion model with measured atmospheric concentration and surface deposition. At the present time, the multi-media environmental pollution caused by the massive release of radionuclides to the atmosphere from the FNPS1 is still severe natural and social issues, while the total amount of source term and its temporal variation has a large uncertainty. In this situation, the author's work brings very valuable and timely information to the international society. The topic of manuscript certainly is suitable for ACP.

➔Thank you so much for your positive comments on our paper toward the publication.

The new source term was validated by comparing the modified WSPEEDI-II simulation using the new source term and the previous WSPEEDI-II simulation using the previous source term with measurements. As a result, both effects due to improved deposition scheme and improved source term are mixed in the discussion for validation of the new source term. This is a weakness in this manuscript. In the analysis, it is very important to separate two effects of deposition scheme and source term. For example, the authors compare the simulations using the new source term and the previous source term based on the modified WSPEEDI-II model and then analysis the differences between two simulations.

➔As suggested by the reviewer, we carried out four cases of simulation in terms of deposition distribution of Cs-137 using the combinations of (1) old source term (Terada et al., 2012) and original WSPEEDI-II, (2) new source term and original WSPEEDI-II, (3) old source term and modified WSPEEDI-II and (4) new source term and modified WSPEEDI-II. By comparing the results from (1) and (3), the effect of model improvement on deposition distribution was examined and then, by comparing (3) with (4), the effect of source term improvement to the distribution was investigated. The results were described in subsection 4.1.1.

Additionally, there are many points which should be clarified. The reviewer recommends publishing this paper with major revisions in response to the following questions and comments.

➔As shown below, we carefully responded to your comments through revision of the manuscript. However, we sometimes could not bring the whole sentences from the manuscript to this letter because they are too long to cite in the letter. We showed the chapter, section, or subsection number, so please also see the revised manuscript.

<u>MAJOR COMMENTS</u>

Validation of new source term Most important progress in author's study is to determine the new and detailed source term of I-131 and Cs-137 combining measurement data and offline coupling model of the atmospheric and oceanic dispersion models. Hence, it is important to compare the modeled results of new and previous source terms using the modified WSPEEDI-II model based on measurement data (air concentration, air dose rate, and surface deposition) and then demonstrate the advantage of the new source term quantitatively. From this viewpoint, the authors need to add analysis and discussion. Additionally, for the air concentration, the authors should use the measurement data at not only JAEA-Tokai but also other sites.

➔ As replied above, we made four cases of simulation in terms of deposition distribution of Cs-137 using the combinations of (1) old source term (Terada et al., 2012) and original WSPEEDI-II, (2) new source term and original WSPEEDI-II, (3) old source term and modified WSPEEDI-II and (4) new source term and modified WSPEEDI-II. By comparing the results from (1) and (3), the effect of model improvement on deposition distribution was examined and then, by comparing (3) with (4), the effect of source term improvement to the distribution was investigated. The results were described in subsection 4.1.1.

Concerning the air concentration data, we did not use them for validation using WSPEEDI-II because these are already used for source term estimation. Though the validation by using some WMO models uses only data of JAEA-Tokai, it is because this data only show the temporal variation in detail and suitable for validation. It may be possible to compare other data, but we hesitate to increase the volume of paper.

Validation of new deposition scheme In this paper, a new deposition scheme, which deals with dry and fogwater depositions, CCN activation and subsequent wet scavenging for radioactive iodine gas and other particles, was incorporated into WSPEEDI-II. However, this new deposition scheme wasn’t validated objectively based on measurement data, though authors discussed about it in section 4.2. The authors should compare the modeled results with old and new schemes in fixed source term and then demonstrate the advantage of the new scheme.

➔ As replied above, we examined the reliability of deposition model and new source term by carefully comparing above four WSPEEDI simulation cases in the revised paper (subsection 4.1.1).

Ratio of gaseous and particulate I-131 The ratio of gaseous and particulate I-131 is one of critical parameters in the inverse estimation for I-131 source term. It is well known that the ratios of gaseous and particulate I-131 have large variability in time and/or space. It is considered that the ratio depends on the source condition as well as the history of air mass (especially, washout or not). In fact, the ratio of gaseous to total I-131 varies in the range of 0.2 to 0.8 in space and/or time (Tsuruta et al., 2012, 2014). The authors should discuss the impacts of the ratio determined from the data collected at only one site (JAEA-Tokai) on the I-131 source term estimation.

➔ As the concentration data of JAEA-Tokai is used for the source term estimation, we also used the ratio of gaseous and particulate I-131 from JAEA-Tokai data. Of course, we consider the temporal variation of the ratio at Tokai and correct the ratio at Tokai to that at the release point considering the history of air mass by simulation model which can treat the difference of deposition process of gaseous and particulate. However, the reviewers comment is important. Thus, we cited (English) paper (Tsuruta et al., 2012) and noted that the ratio depended on the source condition as well as the history of air mass (new subsection 2.3.3).

<u>Individual comments</u>

1) Page 14735, lines 14-15: Why does the use of peak value reduce the impact of discrepancies in arrival time? If the modeled arrival time differs from the measured time, the modeled peak concentration should be different from the measurement. Additionally, the use of peak values, which is not statistically stable, may cause higher uncertainty in the reverse estimation. Why did the authors use peak value, not average value?

➔ The comparison of peak values is carried out for JAEA-Tokai data, because time series data are only available at this place. The air concentration data from JAEA-Tokai is one hour averaged one. When we use the reverse estimation method, "Since the uncertainty of model simulation is the primary cause of the discrepancy in the spatiotemporal distribution of plume between the measurements and simulation results, the above procedures cannot be applied systematically, and the correction of this discrepancy by 'expert judgment' is necessary to reduce the impact of model uncertainty on the source estimation. The process is to check all available measurements to see if the plume is reproduced appropriately or not for comparison with the measurements, and to determine if the discrepancy is caused mainly by errors in the calculated wind direction. If the plume flow direction is clearly different from the measured wind direction, the calculated plume is rotated to match the measured wind direction and Eq. (1) is applied. The use of peak values corrects any discrepancy in the timing of the arrival of the peak air concentrations between the measurement (JAEA-Tokai) and simulation. We assume that the peak values of the measurement and simulation are comparable even though the timing or temporal pattern of the arrival of the peak is different because the central plume axis passes across the sampling point differently between the measurement and simulation" (new subsection 2.1.1).

2) Page 14736, lines 15-16: This part is not clear. The authors should explain more detail.

➔ This part is deleted in the revision to shorten the paper.

3) Page 14737, lines 11-12: The authors should explain how to classify the affected points from the un-affected points.

➔ We revised in subsection 2.1.2 as "Note that by preliminary comparison between measurement points of sea surface concentration of $^{134}$Cs and the oceanic dispersion area estimated by simulation using sources of direct release from FNPS1 into the ocean, only observational points that are not affected by the direct release of $^{134}$Cs from FNPS1 to the ocean are used for the source term estimation.".

4) Page 14739, line 2: It is better that "Fig.2" is changed to "Fig.2d".

➔ Figure 2d was deleted to reduce the paper length.

5) Page 14740, lines 2-4: It is better that Fig. 6 is moved to section 3.1 because this figure is modeled results. Fig. 6 seems to show the modeled surface concentration of plume emitted in the event shown in captions for each panel. It is better that brief explanation of Fig. 6 is added.

➔ (Old) Fig. 6 was deleted according to other reviewer's comment.

6) Page 14742, line 7: According to Fig. 9a, the model extremely underestimates the peak at 15:00. The authors

should add some comments on the reason and effects to source term estimation.

➔ All of source term estimation using air dose rates was carried out by using ground-shine. For example, in (new) Fig. 6a, the timing of peak appearance by the plume arrival, and the values of ground-shine shown as slow-slope after the peak agreed well with observations. Hence, the difference of peak values does not affect the estimation of source term. The usage of ground-shine was addressed in (new) sections 3.1 and 3.4, and subsection 4.1.1.

7) Page 14743, lines 1-3: The modeled narrow deposition band (Fig. 9) shifts around 30 degrees compared with measurement (Fig.10). The authors should add some explanations for the shift.

➔ We added the sentences in section 3.1 as "Figure 6c compares the distribution of air dose rates in day-time of March 13 between the WSPEEDI simulation and observation by portable monitor mentioned above. The calculated result is slightly shifted to the west due to the delay of the wind shift comparing with observed wind shift, but it shows the similar distribution pattern and air dose rates as observed ones."

8) Page 14743, lines 4-5: What is evidence for "deposition area was far from the plant due to the elevated release from the stack"?

➔This sentence was deleted to reduce the paper length. However, the evidence is explained as follows: the radionuclides were released from the stack during the venting event on 12 March. Thus, the downwind distance is required until when the plume disperses and contacted with the ground surface. Since there was no rain event on 12 March, only the dry deposition of the plume formed the contaminated area.

9) Page 14744, line 27: "Fig. 11b" should be changed to "Fig.8b".

➔Thank you for your suggestion. In the revised manuscript, figures are all re-numbered according to organize the contents of the paper.

10) Page 14746, line 9: More detail explanation for "modifications of the deposition scheme" is needed.

➔The sentence was unclear. As explained in (new) subsection 4.1.1, wet deposition of the plume released in the morning of 15 March has been enhanced by high scavenging coefficient of modified deposition scheme. The sentence was now revised in section 3.4: "Due to an increase of wet scavenging coefficient in the modified deposition scheme (Fig. A2), …"

11) Page 14746, lines 10-14: One possible reason for "a large increase in the air dose rates did not appear at Fukushima and Iitate areas" is that most of radionuclides deposited before the air mass arrived at these areas. The authors should analysis and discuss more carefully.

➔As described in section 3.4, we do not think most of radionuclides deposited before the air mass arrived at these areas "because rain bands coming from the northwest during the afternoon of 15 March caused the precipitations started around Iitate area from 16:00, and those were very small about 1 mm $h^{-1}$ (Fig. S1). Moreover, Ohno (4.9 km WSW from the site) had no rain observed until the night (Fig. S1). The fact suggests that the plume discharged in the afternoon should produce less amount of (dry) deposition along the pathway from the FNPS1 to the northwest direction. Therefore, the plume can reach Iitate and increase air dose rate due to wet deposition if a large amount of radionuclides were discharged during the afternoon.”

12) Page 14748, lines 13-23: Both ratio of I-131/Cs-137 and gaseous/particulate iodine on 16 March are higher than those on 15 March. These facts may suggest that particulate Cs-137 and I-131 were deposited by precipitation before arriving at JAEATokai on 16 March while there was no precipitation on 15 March. Hence the authors should analysis and discuss based on the background that the changes in the ratios of I-131/Cs-137 and gaseous/particulate iodine are caused by not only source change but also history of air mass.

➔ We revised the paper in section 3.4 based on your comment. Because the simulation model considers the changes in the ratios of $^{137}Cs/^{131}I$ and of gaseous/particulate iodine caused by transport history of air mass, the ratios at the measurement point is different from those estimated at the release point in this study.

13) Page 14752, lines 19-22: Which is a larger factor in improvement of regional deposition, enhancement of the scavenging coefficient or revised source term?

➔ To respond your major comment, we made several test simulations using original/revised WSPEEDI and Terada/new source terms and evaluate the impacts of modeled deposition scheme and revised source term to obtain important improvements in simulation results. The discussion is available in subsection 4.1.1.

14) Page 14753, line 22: The authors should show the source term for Te-132 and the evidence that the modification for Te-132 is reasonable.

➔ Since (new) Table 6 has no space, we prepared the supplemental .csv file for the new source term for Te-132, Cs-134, with I-131 and Cs-137. Te-132 (its progeny I-132) and I-131 are major nuclides to contribute the air dose rate due to ground-shine in the early stage (“Local depositions of $^{131}I$, and $^{137}Cs$, and air dose rate over Fukushima Prefecture“ in subsection 4.1.1). Since the modified model with new source term well reproduced measured air dose rates (Fig. 11a, d and Fig. 16d), we believe that “the modifications to the deposition scheme and source term are reasonable, particularly for $^{132}Te$ and $^{131}I$, which are the major contributors to the ground-shine in the early phases of the accident.” (subsection 4.1.1)

15) Page 14753, line 27: Table 6 shows that the statistical scores of new results are lower than previous results in air dose rate in the north-west area of FNPS1.

➔As replied above, Table 6 was completely modified for four test simulations using original/revised WSPEEDI and Terada/new source terms to demonstrate the impacts of modeled deposition scheme and the new source term.

16) Page 14755, line 6: Why was the measurement data at only one site (JAEA-Tokai) used?

➔ The validation by using some WMO models uses only data of JAEA-Tokai because this is the only data showing the high temporal variation which is suitable for model validation. Although it may be possible to compare other datasets, we did not use those because we needed to reduce the paper size.

17) Page 14755, lines 10-14: The authors should compare the modeled concentration using the new source term with that using the previous source term in Fig.18 and then demonstrate some advantages of the new source term.

➔ Our discussion of WMO model results confused the reviewer. It was not surprising that the performance of each model in (new) subsection 4.1.2 widely varied because these used the source term derived from the different model (i.e., WSPEEDI-II). We should focus on how the new source term works on atmospheric dispersion simulations to some extent. Thus, we deleted those sentences.

18) Page 14755, lines 22-25: The score of FB in Table 7 shows that the modeled deposition using the new source term tends to underestimate and to be worse than the simulation using the previous source term.

➔ Our discussion of WMO model results confused the reviewer. As explained above, we should focus on how the new source term works on atmospheric dispersion simulations to some extent. Thus, we deleted we extracted all of tables related to this work.

19) Page 14757, lines 5-7: In Table 7, The NMSE and FB for I-131 concentration with the new source term was worse than those with the previous source term excluding FB for MLDP0 model. The authors need to make comment on these results.

➔ As replied above, our statistics table confused the reviewer. It was not surprising that the performance of each model in (new) subsection 4.1.2 widely varied because these used the source term derived from the different model (i.e., WSPEEDI-II). We should focus on how the new source term works on atmospheric dispersion simulations to some extent. Thus, we extracted all of tables related to this work, and focused on the overall agreement between calculations and observations when using the new source term in (new) subsection 4.1.2.

20) Page 14758, lines 4-10: The authors should compare the modeled concentration using the new source term with that using the previous one in Figs. 22 and 23 and then demonstrate the advantage of the new source term.

➔ As replied above, it was not surprising that the performance of each model in (new) subsection 4.1.2 widely varied because these used the source term derived from the different model (i.e., WSPEEDI-II). We should focus on how the new source term works on atmospheric dispersion simulations to some extent. Thus, we extracted all of tables and focused on the overall agreement between calculations and observations when using the new source term in (new) subsection 4.1.2.

21) Page 14758, lines 20-24: This part is not clear and more explanation is needed.

➔ Since this part was confused, we deleted in the revised version.

22) Page 14758, lines 25-28: This is true? According to Table 8, the CCs for particulate I-131 and Cs-137 in the new source term case was slightly lower than that in the previous source term. Additionally, the NMSE became worse when the new source term was used. For other scores, the situation was case by case.

➔ Our statistics table confused the reviewer. As replied above, it was not surprising that the performance of each model in (new) subsection 4.1.2 widely varied because these used the source term derived from the different model

model (i.e., WSPEEDI-II). We should focus on how the new source term works on atmospheric dispersion simulations to some extent. Thus, we extracted all of tables and focused on the overall agreement between calculations and observations when using the new source term in (new) subsection 4.1.2.

23) Page 14761, lines 3-4: From Fig. 26a, it is needed that the "then particulate iodine, and finally gaseous CH3I" is changed to "then gaseous CH3I, and finally particulate iodine".
➔ This context was deleted because we removed section 4.2 of the original version to shorten the manuscript. However, the calculations using the revised WSPEEDI-II model shown in Supplement (Fig. S5) showed that the DRY deposition of the iodine is larger in the order of: Gaseous I2 > Particle I > Gaseous CH3I, which was consistent to what we mentioned in our previous manuscript.

24) Page 14761, lines 10-11: This is true? Fig. 26 shows the gas species of I-131 contribute to the contamination of East Japan though their contributions are lower than wet deposition of particulate I-131.
➔ This context was also removed because we removed section 4.2 of the original version to shorten the manuscript. However, we still corrected the sentence to emphasize the negligible effect of the gaseous iodine on "wet deposition" not "total deposition" in Supplement.

25) Page 14761, lines 18-20: The authors should show the appropriate reference indicating WRF-CMAQ model overestimated the observed precipitation amount over Tochigi and Gunma Prefectures.
➔ According to reviewer comments, we moved the discussion of this part from the maintext to Supplement. In the revised version, this sentence was completely deleted.

26) References: There are some references in review. These references should be changed to alternative references which readers can access.
➔ We added all available references to References section.

28) Fig. 9: The authors should answer the following questions about Fig. 9c. – The time of calculated air dose rate (12:00) is different from the time of measurement (from 6:00 to 15:00). Why did the authors use the data at different time for comparison? The air dose rate has a peak around 15:00 on 12 March as shown in Figs. 9a and 9b. Why did the authors use data in the daytime on 13 March instead of data around 15:00 on 12 March?
➔ According to the simulation, the plume was transported toward the Pacific Ocean on 13 March and, thus, the measured air dose rates in (old) Fig. 9c are due to deposit radionuclides discharged in the afternoon of 12 March. Thus, we compared the calculated air dose rate at 12:00 with measurements from 6:00 to 15:00 on March 13 to validate the source term by comparing the calculated and observed distributions of air dose rates. We added this point in section 3.1.

27) Fig. 8: The color scale bar is invisible and should be improved. In the title of figure, the "spatial distribution of accumulated surface deposition" may be better.

29) Fig. 13: The color lines in each figure are invisible and should be improved to be visible.

30) Fig. 17: The black and red lines and horizontal axis in each figure are invisible and should be improved to be visible.

31) Fig. 18: The vertical and horizontal axes in each figure are invisible and should be improved to be visible.

➔ We revised as suggested by the reviewer. Thank you so much.

## Response to Ref.3

In this paper the authors propose a new estimation of the releases of several radionucleides during the Fukushima accident. For that they use a modified version of the atmospheric dispersion model WSPEEDII – with a new deposition scheme – and the oceanographic dispersion model SEA GEARN-FDM, together with air concentration and surface deposition measurements – augmented with respect to the one used in previous work. Also, a detailed analysis of the recovered source is provided, comparing it with the events that took place during the nuclear accident. A validation of the source is performed by comparing the simulated measurements provided by the new source with the real ones. These simulated measurements are obtained, first, using the WSPEED-II atmospheric dispersion model and then, using other proposed atmospheric dispersion models. The article provides new insights on the releases during the Fukushima accident, and it fits within the scope of Atmospheric Chemistry and Physics. Hence I recommend it for publication after the following comments are addressed.

➔ Thank you for your careful reading and providing critical comments on our paper. We tried to respond to suggestions in detail as much as possible, but we sometimes could not bring the whole sentences from the manuscript to this letter because they are too long to cite in the letter. We showed the chapter, section, or subsection number, so please also see the revised manuscript.

<u>MAJOR COMMENTS</u>

In general, the article is too long. It should be written in a much more concrete and concise way, making it easier to understand for the reader.

➔ On the basis of all reviewer comments, we made the following revisions to reduce the paper length summarized as:

- In Introduction, the sentence after 'First, …' of ln.5/ P.14730 was shorten.
- Chapters 3 and 4 are reconstructed. 3.1.1-3.1.9 which describes the source term and the motion of plume during March was revised and only the periods which were different from the previous paper (Terada et al. 2012) were remained. Related this revision, Figs. 5 and 6 were deleted. As for chapter 4 (Discussion), section 3.2 (Verification of source term) was moved to section 4.1. Section 4.1 (Comparison of source terms) was moved to section 4.2. Finally, section 4.2 (Role in deposition process) was moved to Supplement.
- The section for "Validation using several models" was shortened. Related this revision, (old) Figs. 18, 20, and 22 and Tables 7, 8, and 10 were extracted.
- Conclusion was also shortened based on reviewer's comments.

At the same time, we needed to include more sentences to reflect all reviewers' comments to the manuscript (e.g., Meanwhile, Fig. S3 was moved to the main text as (new) Fig. 12). However, the page numbers and numbers of figures and tables were still reduced compared with our previous paper.

Section 2.2 Why, instead of posing the problem as a linear system (Stohl et al., 2012), the authors estimate the source unknowns one by one? Is there any advantages in using the method proposed in the paper with respect to (Stohl et al., 2012)? Using a linear system, the situation explained in p 14735 | 12 would be solved in a more reasonable way. The same applies for the correction of the source in section 2.3

➔ At the first part of chapter 2, we defined the reverse and inverse methods with merit and demerit and also described why we choose the reverse method for the Fukushima case as "A reverse method evaluates the release rates of radionuclides by comparing measurements of air concentration of a radionuclide or dose rate in the environment with calculated one by atmospheric transport, dispersion and deposition models (ATDM) for a unit release of a radionuclide. The release rate is estimated by the ratio of the measurement to calculation result. The merit of the reverse method is that the comparison can be made with one or more independent data points. For example, the minimum number of data points needed is only one and the measured data used for the estimation can change with time from air concentration to air dose rate and vice versa. The demerit is that this simple comparison without consideration of the uncertainty of the ATDM results may cause the large errors, and, consequently, expert judgment is essential to correct the discrepancy between the measurement and calculation."

The explanation of the estimation methods is, in general, confusing. Many details should be clarified:

How is the temporal discretization of the source defined, i.e., starting and ending points of each temporal element? Why is this discretization not regular? This explanation must be included in the manuscript.

➔ In this study, we divided the release interval and looked for when and where the specific segment of plume increased air dose rate, and then found the appropriate observation data which can be used for the source term estimation for the certain plume. Thus, the timing of releases should be clear. The ground-shine was used for the source estimation on 12 and 15-16 March. For the case of 12 March, the release period for wet venting of Unit 1 was determined from the decrease of pressure of drywell and the release by hydrogen explosion is assumed almost instantaneous. For the case of 15-16 March, the release period was basically divided with every hour and the source term for each plume was estimated by the method as mentioned above. The descriptions how to find the appropriate observation and how to determine the release period of the segment of plume are available in (new) subsections 2.1.1 and 2.3.3.

The mathematical notation is, in general, quite confusing and makes the method description unnecessarily difficult to understand.

Eq. (1), (3): Qi, Mi and Cidepend not only on space, but also on time. This must be indicated (Qi(t), for example). Also it is necessary to make clear the difference between time of emission and time of detection. For example, in Eq.1, the time of Qi and time of Mi are different, Qi(t1) , Mi(t2).

➔ We modified the equations and these explanations in subsections 2.1.1 and 2.1.2.

Section 2.2 and 2.3 The measurements and the dilution factors may contain errors. The estimated source may be sensitive to these errors. How do you address this problem?

➔ We realize the importance of the evaluation of the reliability and uncertainties. However, the detailed discussion on this matter needs more space in paper. Thus, we would like to discuss this matter shortly in Conclusion. Detailed discussion will be considered in future paper with discussion of the effect of our deposition model.

p 14735 | 11 Explain why only the peak values are used.

➔ The comparison of peak values is carried out for JAEA-Tokai data, because time series data are only available at this place. The air concentration data from JAEA-Tokai is one hour averaged one. When we use the reverse estimation method, "Since the uncertainty of model simulation is the primary cause of the discrepancy in the spatiotemporal distribution of plume between the measurements and simulation results, the above procedures cannot be applied systematically, and the correction of this discrepancy by 'expert judgment' is necessary to reduce the impact of model uncertainty on the source estimation. The process is to check all available measurements to see if the plume is reproduced appropriately or not for comparison with the measurements, and to determine if the discrepancy is caused mainly by errors in the calculated wind direction. If the plume flow direction is clearly different from the measured wind direction, the calculated plume is rotated to match the measured wind direction and Eq. (1) is applied. The use of peak values corrects any discrepancy in the timing of the arrival of the peak air concentrations between the measurement (JAEA-Tokai) and simulation. We assume that the peak values of the measurement and simulation are comparable even though the timing or temporal pattern of the arrival of the peak is different because the central plume axis passes across the sampling point differently between the measurement and simulation" (new subsection 2.1.1).

p 14736 | 25 How do you determine in which periods the plume flows towards the ocean?

➔ This was written in (new) subsection 2.1.2 as "The judgment of whether the plume during each segment directly flowed toward the ocean is done by evaluating the simulation of the modified WSPEEDI-II, observed wind direction, and monitoring data on the land."

Section 2.3 In general, the subindexes of the variables are extremely confusing here, because they mix space and time. To make clearer what is what, the notation must be revised completely. For example in Eq. (4), Cnj [k], instead of Cj,k

➔ The suffix and equations in (new) subsection 2.1.2 are revised according to the comments from each reviewer.

p 14738 | 2 Where does this equality come from? A more detailed explanation should be included.

➔We revised the related sentences in (new) subsection 2.1.2.

p 14760 | 14 In 14759 | 13, you use the events that took place during the accident to assess your source, and thus claim that your source estimation is correct, and previous source estimations in the literature are not correct. But later, you compare again to the same previous estimations in the literature (which are supposedly wrong) and where they agree with your results, you claim that this again confirms the correctness of your results. This is not a consistent argument!

➔ At page 14759 line 13 in our previous paper, it is described that the source term in the afternoon to the night of 15 March in present study is better than previous one, while at page 14760 line 14, the estimations in present and previous studies when the plume flowed toward the ocean are resemble. The evaluated period is different.

Minor comments

p 14735 | 10 What does it mean "if the data show a continuous time series"? The measurements are always discrete.
➔We deleted this sentence in the revised version.

p 14737 | 15 The variable name Cos can be mistaken with a cosine. A different name should be used.
p 14737 | 11 typo "Note that the only the observational. . ."
p 14738 | 2 Equation number missing.
➔ We modified the above mistakes. Thank you for your suggestions.

Section 3.2.1 Does it have sense to validate the source with the same measurements and model that were used to estimated it? Because the source that fits best with these measurements is the one that produces overfitting. If the same model and the same measurements are used for validation, then some kind of cross-validation technique must be used.
➔ As described in section 2.2.3, for validation, we did not use the same measurements used to estimated source term but used surface deposition, air dose rate, and global air concentration data.

Section 4.1 The wet venting at Unit 3 and DW pressure deficits do not directly imply that the major release took place at this time. If it does, you should argument it properly.
➔ As you suggested, it is not sure that the venting at Unit 3 generated the MAJOR release or not. However, we believe that it is reasonable to assume that DW pressure deficits imply the atmospheric releases. In the case of the evening to night of 15 March, we cannot determine the ratio of releases from Units 2 and 3 due to lack of information. Thus, though we assumed both are the source of releases we changed the sentences less conclusive as "As a result, the period of the potential major release is coincident with the wet venting at Unit 3 and/or DW pressure deficits at both Units 2 and 3 reported on 15–16 March (Fig. 9)." in section 4.2.

p 14760 | 1 How do you know that the source changes drastically in this period? How do you asses that?
➔ From environmental monitoring data, it was expected that a large amount of releases occurred from 15-16 March. However, the reason is still not clear. Thus, the source term is estimated in detail which is one of the main objectives of our paper. We added this point in subsection 2.3.3.

A general correction of minor typos through the whole paper is necessary.
➔ We checked and revised typos throughout the manuscript (including English corrections).

## Response to Ref.4

This paper presents a very detailed analysis of the complex release and dispersion situation for the Fukushima reactor accident performed mainly by the Japan Atomic Energy Agency JAEA. It provides both a better methodological background and an update of the preliminary source term of Chino et al. (2011) which has found widespread use. Therefore, this work is a useful addition to the body of literature on the meteorological aspects of the accident.

➔ We appreciate your crucial comments and suggestions on our paper. We tried to respond to suggestions in detail as much as possible, but we sometimes could not bring the whole sentences from the manuscript to this letter because they are too long to cite in the letter. We showed the chapter, section, or subsection number, so please also see the revised manuscript.

<u>MAJOR COMMENTS</u>

1. The paper is very long and not easy to read, even though certain important aspects are still not sufficiently covered. I would recommend that the paper is revised in a way that it would be more systematic, focus on the most important aspects, and would defer minor aspects to the Supplement. For example, the discussion of the single phases on different days of the accident could be trimmed down, moving a systematic description to the Supplement. However, this description should then be really systematic, best in the form of a table with standard information for each phase, possibly also related figures.

On the basis of all reviewer comments, we made the following revisions to reduce the paper length summarized as:

- In Introduction, the sentence after ‘First, …’ of ln.5/ P.14730 was shorten.
- Chapters 3 and 4 are reconstructed. 3.1.1-3.1.9 which describes the source term and the motion of plume during March was revised and only the periods which were different from the previous paper (Terada et al. 2012) were remained. Related this revision, Figs. 5 and 6 were deleted. As for chapter 4 (Discussion), section 3.2 (Verification of source term) was moved to section 4.1. Section 4.1 (Comparison of source terms) was moved to section 4.2. Finally, section 4.2 (Role in deposition process) was moved to Supplement.
- The section for “Validation using several models” was shortened. Related this revision, (old) Figs. 18, 20, and 22 and Tables 7, 8, and 10 were extracted.
- Conclusion was also shortened based on reviewer’s comments.

At the same time, we needed to include more sentences to reflect all reviewers’ comments to the manuscript (e.g., Meanwhile, Fig. S3 was moved to the main text as (new) Fig. 12). However, the page numbers and numbers of figures and tables were still reduced compared with our previous paper.

The authors give now more insight into their method of source determination. This is really important as previous publications have not been very explicit on this topic. However, the presentation should still be improved, and be better placed in the context other similar work, as is detailed below.

2. The authors use both the terms “reverse estimation” and “inverse estimation”. It seems that the latter is reserved for the part of the source reconstruction using concentration data in ocean water. The authors should explain what

they mean by these terms and consider established technical terms. In applied mathematics and related sciences, the term “inverse problem” and derived from it “inverse method / modelling / . . . ” is the standard language (however, as will be explained below, the method does not correspond to a formal inversion).

➔ According to the reviewer comment, we defined the reverse and inverse methods with merit and demerit and also described why we choose the reverse method for the Fukushima case at the first part of chapter 2 as “A reverse method evaluates the release rates of radionuclides by comparing measurements of air concentration of a radionuclide or dose rate in the environment with calculated one by atmospheric transport, dispersion and deposition models (ATDM) for a unit release of a radionuclide. The release rate is estimated by the ratio of the measurement to calculation result. The merit of the reverse method is that the comparison can be made with one or more independent data points. For example, the minimum number of data points needed is only one and the measured data used for the estimation can change with time from air concentration to air dose rate and vice versa. The demerit is that this simple comparison without consideration of the uncertainty of the ATDM results may cause the large errors, and, consequently, expert judgment is essential to correct the discrepancy between the measurement and calculation.” Also, we changed the description for the method used to assess the source term over the ocean from “inverse” to “reverse”.

3. The authors explain that for the land data, they proceed as follows, with my questions in brackets:

(a) Divide the time axis into intervals [Obviously, this is largely subjective. What is the role in this division of steadiness of meteorological conditions, steadiness of the plant state, steadiness and/or availability of measurements?]

➔ In this study, we divided the release interval and looked for when and where the specific segment of plume increased air dose rate, and then found the appropriate observation data which can be used for the source term estimation for the certain plume. Thus, the timing of releases should be clear. The ground-shine was used for the source estimation on 12 and 15-16 March. For the case of 12 March, the release period for wet venting of Unit 1 was determined from the decrease of pressure of drywell and the release by hydrogen explosion is assumed almost instantaneous. For the case of 15-16 March, the release period was basically divided with every hour and the source term for each plume was estimated by the method as mentioned above. The descriptions how to find the appropriate observation and how to determine the release period of the segment of plume are available in (new) subsections 2.1.1 and 2.3.3.

(b) Only a single measurement available for the respective release period: pair maximum of both measurement and model to determine source. [Does it really occur often that only a single measurement is affected? This seems unlikely. Deposition data are always available, even though they were not used here. And how can one ascertain that only a single release phase impacts this measurement? What if this is not the case? What is the time and space window considered for matching the maxima? Even though this method should increase the robustness of the result, it could also be a major source of error if something is paired which does not match.]

➔ Yes, the dust sampling in Fukushima Prefecture was carried out by limited number of monitoring cars and consequently, the data are not temporally successive and the number of data was small per day. Exceptionally, at

JAEA-Tokai, temporal variation of air concentrations was observed continuously (subsection 2.2.1). Thus, the number of measurement points for air concentration which are the most suitable for the source estimation was small, and furthermore the data which successfully measured the concentration of plume was almost one or two in the same time. The number of available data can be found by looking for when and where the specific plume increases air dose rate by the WSPEEDI simulation (subsection 2.1.1). Unless we get the temporal variation of deposition, it cannot be used for the reverse estimation. We do not intend to estimate the source term for each event in each reactor. As you know, the source term estimated by reverse estimation is always the total release rate from three reactors. However, when the peak position in both measurement and calculation which assumed the release of specific event is the almost same, it is natural to think that the measurement values is mainly due to the specific event. The window of errors between calculation and measurement are up to three hours at JAEA-Tokai.

(c) If multiple measurements are affected, average both measurements and model. [Will again only the peaks be considered, or will whole time series over some interval be averaged? What is the justification of this approach – should one not rather average the resulting ratios?]

➔ The comparison of peak values is carried out for JAEA-Tokai data, because time series data are only available at this place. The air concentration data from JAEA-Tokai is one hour averaged one. When we use the reverse estimation method, "Since the uncertainty of model simulation is the primary cause of the discrepancy in the spatiotemporal distribution of plume between the measurements and simulation results, the above procedures cannot be applied systematically, and the correction of this discrepancy by 'expert judgment' is necessary to reduce the impact of model uncertainty on the source estimation. The process is to check all available measurements to see if the plume is reproduced appropriately or not for comparison with the measurements, and to determine if the discrepancy is caused mainly by errors in the calculated wind direction. If the plume flow direction is clearly different from the measured wind direction, the calculated plume is rotated to match the measured wind direction and Eq. (1) is applied. The use of peak values corrects any discrepancy in the timing of the arrival of the peak air concentrations between the measurement (JAEA-Tokai) and simulation. We assume that the peak values of the measurement and simulation are comparable even though the timing or temporal pattern of the arrival of the peak is different because the central plume axis passes across the sampling point differently between the measurement and simulation" (new subsection 2.1.1).

4. Usage of dose rates: is it true that only 5 nuclides are considered? Which fraction of the total dose rate can they explain? How does this fraction depend on time and distance?

➔ We estimated the importance of other nuclides for air dose estimation, and recognized it was needed to consider I-133. Thus, we re-estimated the source term considering the release of I-133 in the period using air dose rate (ground-shine) data. These points are added to subsection 2.3.2.

5. The paper is quite obscure concerning how decay is considered, which is relevant for iodine and other short-lived nuclides. This pertains for example to Figure 2 and specifically 2c as well as the adopted I-131/Cs-137 ratio. It must be clearly stated, for all the data (measured, modelled, release) whether they are decay-corrected (and to

which time) or not. If not, how can the ratio be assumed constant over days? Is the ratio only used when the source estimate is based on dose rates, or throughout? I tried to analyse this nuclide ratio from the release data in the supplement and found that after ca. 8 days, it suddenly drops from 10 to 5. It may be just coincidental, but this looks like decay has been ignored for 8 days, then some I-131 measurements come in and the ratio drops corresponding to the decay. Soon after, the ratio jumps by a factor of almost 100 within very short time. Unless a convincing explanation is given for that, I don't think that these ratios and thus the iodine releases can be considered reliable, even though I admit that the comparison with the airborne survey is a good support for the source term. However, as also the deposition parameterization has been strongly modified, and as the reconstructed source term has such a large temporal variability, this evidence is not totally conclusive. I am also surprised by the large scatter of the ratios presented in Figure 2, they need a proper explanation, including uncertainties of the measurements.

➔ Our idea for radioactive decay was described in (new) subsection 2.3.3 as "The ratio of $^{137}$Cs to $^{131}$I at the released time should be different from that at the measurement time because of radioactive decay during the transport of the plume and the difference of deposition processes of both nuclides in the environment. However, the transport time period between the FNPS1 and the monitoring points used to determine $^{137}$Cs/$^{131}$I ratio (Fig. 4c) are within about 10 hours and sufficiently small compared with decay constants of both nuclides (Table 5). Thus, we only considered the latter effect to adjust the ratio obtained at the measurement points to that at the release point." Concerning the $^{137}$Cs/$^{131}$I ratio, we can also estimate the ratio of inventory in the reactors with time considering the decay constants of each nuclide. However, to estimate the ratio when discharged into the atmosphere, it is the only way to believe the ratio measured in the environment even though the ratio highly varies with time for transport and deposition processes. We are also surprised at such large variations in observational data (old Fig. 2), but we do not have further information to clarify the reason, such as the error of measurement and reactor conditions related to the volatility of each radionuclide.

6. The releases tabulated in the supplement contain three physical species of iodine. It should be explained in the paper how they were derived, which measurements were available, and what the uncertainties are for this subdivision.

➔ The ratio of gaseous and particulate iodine is determined from the measurement at JAEA-Tokai. Regarding iodines, "because there are no observed data on the ratio of elemental iodine ($I_2$) and organic iodine ($CH_3I$), the ratio of gaseous $CH_3I$ to $I_2$ is assumed to be constant of 0.6 throughout the simulation period according to the method of RASCAL 4.0 (US NRC, 2012)" (subsection 2.3.3). As written in (new) Conclusion, the ratio "can cause the errors in estimated results affected by wet deposition (Fig. S5b)". The estimation of source term for iodine is also affected by this uncertainty.

7. The explanation in Section 2.3 is difficult to follow. It should be more formalized, but with proper notation. It is not helpful how variable names Ci, Co, Mo are used. It would be better to use a standard notation such as c for concentration, with a subscript such as, e.g., m and o for model and observation, and not to use the letter C for something which isn't concentration – better call the correction factor r or f or similar. Also, if the index j is

occurring everywhere, it can just be dropped. Why is a log function used for averaging? Also, some variables are written upright, others in italics, etc.

➔ We modified the equations and these explanations in subsections 2.1.1 and 2.1.2.

8. As visible in the supplement, various assumptions for the source geometry are made for different release periods. Some explanation on that is needed in the methods sections.

➔ We added the detailed explanation to (new) subsection 2.3.2 (Source assumption).

9. A discussion of the method for source estimation is required. It should include reference to less subjective, formal methods for solving inverse problems related to atmospheric dispersion which are available and partly have already been applied to the Fukushima accident. It should then be explain why the authors believe that their own method, which requires a lot of subjective decisions and manual intervention, is preferable. I do understand that in this situation, very complex both with respect to the source processes and the meteorological phenomena, with data of different type and degrees of quality, a selection guided by all the available knowledge can be very useful in obtaining a robust result, and agencies tasked with this emergency have certainly accumulated a lot of such knowledge. On the other hand, there is also the risk of mistakes in the process and of underusing available information through the lack of a comprehensive method. Personally, I would believe that it should also be possible to use a more comprehensive, formal inverse method together with such knowledge. In any case, the impression should be avoided that the approach used here would be the best or only one for dealing with unknown source terms in a nuclear emergency. It would be good if the authors would frankly explain their experiences and rationale for their approach, and to document their knowledge so that others can also use it (see my suggestion above, for a systematic coverage of the release in the Supplement).

➔ At the first part of chapter 2, we defined the reverse and inverse methods with merit and demerit and also described why we choose the reverse method for the Fukushima case as “A reverse method evaluates the release rates of radionuclides by comparing measurements of air concentration of a radionuclide or dose rate in the environment with calculated one by atmospheric transport, dispersion and deposition models (ATDM) for a unit release of a radionuclide. The release rate is estimated by the ratio of the measurement to calculation result. The merit of the reverse method is that the comparison can be made with one or more independent data points. For example, the minimum number of data points needed is only one and the measured data used for the estimation can change with time from air concentration to air dose rate and vice versa. The demerit is that this simple comparison without consideration of the uncertainty of the ATDM results may cause the large errors, and, consequently, expert judgment is essential to correct the discrepancy between the measurement and calculation.” Also, we changed the description for the method used to assess the source term over the ocean from “inverse” to “reverse”.

10. The comparisons based on the regional-scale WMO calculations show, all in all, lower scores when the new source term is used compared to Terada 2012. Interestingly, NAME with ECMWF data performed best, even though ECMWF is said to have precipitation not well reproduced in this case. There is no discussion of these findings, however, they may indicate a kind of overfitting – or were WMO models tuned to the Terada 2012 source

term?

➔ Our statistics table using WMO models confused the reviewer. There was no tuning to WMO models for Terada et al. (2012) as “To evaluate the new source term independently of the one dispersion model used to develop the source term, numerical simulations from three atmospheric dispersion models were compared to observations using our new source term estimates” in subsection 4.1.2. Thus, it was not surprising that the performance of each model in (new) subsection 4.1.2 widely varied because these used the source term derived from the different model (WSPEEDI). We should focus on how the new source term works on atmospheric dispersion simulations to some extent. Thus, we extracted all of tables appeared in (new) subsection 4.1.2 and focused on the overall agreement between calculations and observations when using the new source term as “Then, the new source term is further tested using different atmospheric dispersion and meteorological models over regional- and global-scales to evaluate its reliability for general atmospheric dispersion model studies during the FNPS1 accident.” in section 4.1. By the way, the discussion of NAME with different meteorological fields was extracted from the maintext to reduce the manuscript length.

11. Global HYSPLIT simulations and comparison with measurements: I consider this part a candidate for removal from this paper. Global simulations and comparisons with distant sites are not that useful for verifying the fine details of the source with which the present paper is mainly concerned with. There are some aspects of the results that would need to be addressed in more detail, such as the underprediction of particulate iodine and the lack of correlation for caesium. Note that the authors speculate that the correlation for iodine is caused only by the co-factor of age and associated decay – see my comments above on decay correction! I was also wondering why the time series plots are almost all clipped so that they don’t show the arrival of the plume properly. Better work on a separate study than reporting preliminary results without proper explanation.

➔We believe that showing the global HYSPLIT calculation result is valuable to verify the new source term in particular when the plume flowed directory toward the Pacific Ocean estimated in this study. However, since the discussion should be limited for reducing the manuscript length, we deleted several figures and focused on only the first arrival time of the plume and the general time series (old Figs. 5, 6, and 22). Further evaluation is needed in future.

For your other questions: radioactive decay is considered in the WSPEEDI model as mentioned before. The time series plots in all figures related to global HYSPLIT calculations have included all available data from CTBTO from the beginning of FNPS1 accident.

12. Discussion of source terms (Section 4.1): I don’t feel so much convinced with these arguments. The increased wet scavenging might be too high, and estimates of dose rates at nearby monitoring stations would depend on effective source heights, which are not so well known and which appear to have been set to some predetermined values based on expert judgement only (the paper is not explicit on this). The emissions are partly attributed, for example, to wet venting of unit 3, “wet venting” or “dry venting” is also referred to in other places. Even though the emergency staff tried to conduct such operations, often it is also reported that “success is not clear” or that operations failed. Furthermore, it is still not well known which release paths to the environment were created

during the course of the accident. Thus, one should not be certain that venting operations reported represent properly the release paths. All in all, results therefore may be not as robust as they are appear in the paper.

➔ As suggested by the reviewer, our description in section 4.1 was too conclusive despite many uncertainties. Thus, we addressed the possible reasons which may cause the uncertainty of new source term in section 4.1 and Conclusion.

13. Supplement with release rates: when I open the xlsx file, I am told that a link to an external file exists (which of course is not present). In view of long-term archival, I don't think xlsx is a good format to include the data – a simple text file could be more useful (including both might be an attractive option). Also, give the units for all the values (it is not stated that releases are given as Bq/h). Do not mix text and data, or source geometry that cannot be read automatically, but provide a format allowing data to be easily ingested by programmes.

➔ We changed the format from .xlsx to .csv. The revised file includes the unit of Bq $h^{-1}$.

14. Measurement data: The paper refers to various sources of measurements, usually web resources. However, many are in Japanese, and usually in the form of PDF files. This kind of sources is not very suitable for use in further studies. I would like to encourage the authors to do their best to contribute to a collection of the relevant environmental data in an accessible, machine-readable format, and provide the most useful sources in their references.

➔ We realize the importance to make the database for monitoring data in English which can be widely used to public. However, at the current stage to respond to all reviewer comments, we have no time to do so and are also afraid that we take mistakes in alternation.

<u>Other comments</u>

1. I think Eq. 2 is trivial enough to be skipped (and writing whole phrases as subscripts is not very appropriate).

➔ The equation was removed.

2. Reference CTBTO (2011) is missing.

➔ The reference is now available in References section as "CTBTO (Comprehensive Nuclear-Test-Ban Treaty Organization, Preparatory Commission): Fukushima-related Measurements by CTBTO, available at: http://www.ctbto.org/press-centre/highlights/2011/fukushima-related-measurements-by-the-ctbto (last access: 25 April 2014), 2011."

3. Figures 12 and 15 should match better (same order of variables shown, domain)

➔The figures are unified to the same figure (new Fig. 11) in the revised version with the same domains and the same order of variables.

4. Figure 25: It would be better to separate this important overview into a Cs-137 and I-131 part.

➔ The figure was separated to (new) Figs. 22 and 23 for Cs-137 and I-131, respectively.

5. Figure 17: It is much too small. Showing the Cs-137 results is sufficient, the others are very similar.
➔ As suggested by the reviewer, we showed the figure for only Cs-137 concentration in Fig. 17, resulting in producing larger graphs.

6. I found Figure S3 quite interesting and was wondering why this Figure was moved to the supplement. Concerning the way it is referenced in the text, this type of figure is not suitable to show that for the majority of monitoring sites the model was within a factor of 2 of the observation. There is a number of sites with nice agreement, but also sites which don’t agree and where the new version did not improve much. It would also be of interest to know whether the sites are a subset, or include available ones.
➔ According to your comment, (old) Fig. S3 was moved to subsection 4.1.1 and reffered as (new) Fig. 12 to demonstrate the impact of the new source term from 15-16 March.

7. Figure 18: Not needed and unreadable. Just give a score such as FA5 as a small table.
➔To make the paper shortened, we deleted this figure in the revised version.

8. Figure 19: Why is the color scale totally different than in the observation plot (Fig. 12)? Also it is very difficult to find geographical features in both figures for comparison. Remove the province border lines and show a well-readable geographic grid, same in both figures. Restrict the area in this Figure to the one in Figs. 12 and 15.
➔ According to the suggestion, we unified the color scale for all figures of spatial distribution so that the reader can compare each other (new Figs. 11, 13, 14, 15, 18, S3, andS9).

9. Tables 6–8: This presentation is not sufficiently clear. All comparisons should be done with the same metrics and be shown in the same layout (a clear one!). My preference is FA2/FA5/FA10 (maybe not all of them are needed), correlation coefficient and bias. FAx is the best metric for these quasi-lognormal data. Bias should not be normalized by mean of obs and model (FB) – gives better score to overpredicting model. As all comparisons in one class have the same obs data, no need for normalisation. Or if you want, normalise with the mean observation only. Add units for all columns which contain dimensional data.
➔ To make the paper shortened, we deleted Tables 7-8 for WMO model calculations in the revised version. In addition, Table 6 has been fully revised with the statistics of CC, NMSE, FB, and FA2/FA5/FA10 to demonstrate the difference between four WSPEEDI simulation cases (combinations of the original/modified WSPEEDI and Terada/new source terms) in subsection 4.1.1.

10. Many of the figures (also some tables) are to small to be legible when printed, only on the screen with 200-400% magnification they can we well read (but some are blurred at high resolution, even worse). Please make sure in the final version that all figures included are readable in print format.
➔We adjusted the figure sizes so that the readers can see figures readily.

11. Several captions of figures with model–observation comparisons lack information about the observation data set used.

➔ We added the references to figures from Figs. 11-16.

12. I don't think having an appendix and a supplement is needed. I would suggest to move the description of the deposition schemes to the supplement.

➔ The description for deposition schemes is necessary for readers to know the changes from the previous version of WSPEEDI-II (particularly in subsection 4.1.1). Therefore, we remained the part in Appendix section with the short explanation of the below-cloud scavenging to answer other reviewers.

13. Language proofreading would be a plus, especially as native English speakers are on the list of authors. (Just a single observation: don't use the word 'trend' for general temporal variation.)

➔ One of our co-authors corrected the linguistic problem throughout the manuscript.