The dataset is potentially useful, but the manuscript does not sufficiently demonstrate its added value over existing satellite-based gap-free global soil moisture datasets. The validation is extensive in quantity, but much of it repeatedly compares the final product with other soil moisture datasets or reconstruction methods. Additional evaluation of physically and meteorologically meaningful soil moisture variability is needed before the dataset can be considered reliable for broader Earth system applications.

Re: Thank you very much for your positive comments on our paper. We have carefully considered your comments and revised the paper accordingly. Please see the response below.

Major comments

1. Clarify the novelty and added value of the dataset. As the authors state in the introduction, many satellite-based global soil moisture datasets already exist. In addition, a previous gap-free FY-3B soil moisture dataset has already been developed. Therefore, the authors should clearly explain what new scientific or practical benefit this dataset provides compared with existing global soil moisture products and the previous FY-3B reconstruction dataset.

Re: Thank you very much for this important comment. We will further clarify the methodological advancement and practical value of the proposed model and dataset and add a comparison experiment with existing global SM products. First, compared with the previously developed spatially seamless FY-3B SM dataset based on the TTP model, GSP represents a further methodological advancement by explicitly exploiting both temporal and spatial neighborhood information. While TTP mainly relied on the temporal information of the target pixel, GSP additionally utilizes valid observations in the surrounding spatial neighborhood, allowing more effective use of the spatiotemporal information in FY-3B SM observations. This methodological advancement will be described in greater detail in the revised manuscript. The performance of GSP over TTP is further supported by the validation results in Section 5.2, including both *in-situ* site-level validation and triple-collocation (TC) validation. Second, the proposed dataset provides a spatially seamless FY-3B-based SM data source that complements existing global SM products. As relatively few studies have focused on reconstructing FY-3B SM into a spatially seamless dataset, the proposed dataset can improve the usability of FY-3B SM observations for various applications. In the revised manuscript, we will further strengthen the discussion of its novelty and value and, to further evaluate its effectiveness, add a comparison experiment with other existing global SM products.

2. Limited physical and meteorological validation. The reconstruction is based almost entirely on FY-3B soil moisture itself, without using meteorological forcings or ancillary variables. Also, the use of a four-month temporal mean in the feature fusion step may smooth the reconstructed values toward the average state. The authors should test whether the final dataset can capture physically meaningful soil moisture variability, such as seasonal variations, wet and dry events, and responses to rainfall/temperature changes.

Re: Thank you very much for this valuable comment. We agree that the physical and meteorological consistency of the reconstructed FY-3B SM dataset should be further evaluated. To address this concern, we will add additional analyses in the revised manuscript. First, to evaluate whether the reconstructed FY-3B SM can capture physically meaningful responses to precipitation events, we will add precipitation variations in Section 4.1 and compare them with the temporal variations of the reconstructed FY-3B SM. We will also compare the reconstructed SM time series with *in-situ* data to further evaluate its ability to capture seasonal variations and short-term SM dynamics. Second, to examine whether the four-month temporal mean used in the feature fusion step causes the reconstructed SM to become overly smoothed toward the mean state, we will further analyze its temporal variability and anomaly characteristics. Specifically, we will compare the temporal standard deviations of the reconstructed and *in-situ* data, and evaluate the anomaly correlation between the reconstructed and *in-situ* data. These analyses will help determine whether the reconstruction preserves meaningful temporal variability and wet/dry anomalies.

3. Need for uncertainty information and clearer data limitations. The validation primarily relies on artificial masked FY-3B observations. While this approach is useful, it does not fully prove the accuracy of reconstructed values under real missing-data conditions (e.g. frozen soil, snow, vegetation effects). The final dataset should include uncertainty or quality information, especially to distinguish direct satellite retrievals from reconstructed estimates. Based on this information, the authors should also provide clear guidance for users on data reliability and potential limitations.

Re: Thank you for this important comment. In the revised manuscript, we have therefore further clarified the complementary roles of the simulated-gap and *in-situ* validation experiments, added source information to distinguish original and reconstructed observations, provided information on reconstruction uncertainty, and expanded the discussion of the potential limitations of the dataset.

First, regarding that the validation primarily relies on artificial masking, we would like to clarify that the simulated-gap experiments were designed to reproduce realistic missing-data conditions rather than simply generate random artificial gaps. Specifically, the masks used in the simulated-gap experiments were derived from the actual missing-data patterns corresponding to the testing period observed in the FY-3B SM dataset. These masks preserve the spatial characteristics and distribution patterns of real observational gaps. We have added more details in Page 6, Lines 241-245 and Page 11, Lines 383-386. Moreover, the simulated-gap experiments can characterize the empirical reconstruction uncertainty under different missing-data conditions. The RMSE were 0.0119, 0.0119, 0.0121, and 0.0121 m^3/m^3 for missing rates of 30–40%, 40–50%, 50–60%, and 60–70%, respectively. These results indicate that the reconstruction error generally increases as the missing rate increases. Therefore, these statistics can be used as references for the expected reconstruction error and data reliability under different missing-data conditions. We emphasize that these values are empirical error estimates derived from simulated-gap experiments and should not be interpreted as strict pixel-level probabilistic uncertainty bounds. Based on these results, we have also added clearer data-use guidance to the Readme.txt and Page 17, Lines 467-469. Users are advised to first distinguish original and reconstructed values using the source flag and to consider the corresponding empirical reconstruction error when using reconstructed estimates.

Second, the network-level *in-situ* validation complements the simulated-gap experiments by providing an independent evaluation under actual missing-data conditions. Specifically, the in-situ observations were used to evaluate reconstructed FY-3B SM at locations corresponding to actual missing satellite observations, rather than artificially masking FY-3B SM. We also acknowledge that even the *in-situ* validation cannot completely isolate the uncertainties associated with individual physical causes of satellite observation gaps, such as snow cover, frozen soil, and dense vegetation. These details have been explicitly added in Page 12, Lines 399-404.

Third, to improve the quality of the final dataset and clearly distinguish direct satellite retrievals from reconstructed estimates, we have added a source flag dataset. This allows users to directly distinguish observations from reconstructed values and prevents the two types of data from being treated as equivalent in subsequent applications. The final dataset has been updated accordingly, and the corresponding description has also been added to the Readme.txt.

Finally, we have expanded the discussion of the potential limitations of the GSP-based reconstruction framework in the revised manuscript. Since GSP relies entirely on the available spatiotemporal information within the FY-3B SM dataset for reconstruction, the availability and representativeness of valid neighboring observations have a substantial influence on reconstruction performance. When missing observations become extensive or spatially/temporally continuous, the amount of effective neighboring information available to constrain the reconstruction is reduced, which may lead to increased reconstruction uncertainty and degraded accuracy. We have explicitly added this limitation to the Page 28, Lines 627-633.

Page 6, Lines 241-245

"For the training dataset, the global FY-3B SM masks from 16 July 2011 to 31 December 2016 are traversed, and mask patches (with the spatial size of 40×40) with missing rates between 30 % and 70 % are selected. For the testing dataset, an independent mask set is constructed from the original FY-3B SM missing patterns from 1 January 2017 to 15 August 2019, using the same selection criteria."

Page 11, Lines 383-386

"Similar to the simulated mask process during training, the real missing patterns from 1 January 2017 to 15 August 2019 in FY-3B SM were applied to simulate partial spatial gaps to evaluate the GSP performance under different proportions of spatial missing (i.e., 30 %-70 %)."

Readme.txt

"## Source flag

A source flag is provided to distinguish background, original, and reconstructed pixels.

0: Background pixels

1: Original FY-3B SM observations

2: Reconstructed FY-3B SM values

Users are advised to use the source flag to distinguish original observations from reconstructed estimates when using the dataset.

Reconstruction uncertainty and data-use guidance

The reconstruction uncertainty was empirically evaluated using simulated-gap experiments based on actual missing-data patterns observed in FY-3B SM.

The RMSE values were 0.0119 m3/m3, 0.0119 m3/m3, 0.0121 m3/m3, and 0.0121 m3/m3 for missing rates of 30–40%, 40–50%, 50–60%, and 60–70%, respectively.
These values provide empirical references for the expected reconstruction error under different missing-data conditions.
The empirical RMSE values should not be interpreted as strict pixel-level probabilistic uncertainty bounds.
When using reconstructed SM estimates (source flag = 2), users are advised to consider the corresponding empirical reconstruction error associated with the missing-rate range and to distinguish reconstructed estimates from original observations (source flag = 1) using the source flag."

Page 17, Lines 467-469

"The accuracy metrics of GSP in simulated-gap experiments further provide the uncertainty references when using reconstructed FY-3B SM under the corresponding missing-data conditions."

Page 12, Lines 399-404

"To further evaluate the reconstructed FY-3B SM datasets under actual missing-data conditions, two types of in-situ networks, including sparse and dense networks, were employed. The in-situ observations (i.e., 17 sparse networks and five dense networks) were evaluated at locations and dates corresponding to actual missing FY-3B SM observations. Therefore, this network-level in-situ validation complements the simulated-gap experiments by providing an independent assessment of reconstruction performance under actual missing-data conditions."

Page 28, Lines 627-633

"Firstly, the proposed GSP framework has a limitation in that its reconstruction relies entirely on the available spatiotemporal information within the FY-3B SM dataset. When the effective neighboring information is substantially reduced, the reconstruction uncertainty may increase. To alleviate this limitation, future studies could consider incorporating heterogeneous auxiliary data, such as precipitation and LST, to provide complementary information and enhance the robustness of GSP under large-scale and continuous missing-data conditions."

Minor comments

1) Fig. 1: The definition of "temporal coverage proportion" should be clearly explained. How can some regions reach 100% coverage for a LEO satellite?

Re: Thank you for this comment. We sincerely thank the reviewer for pointing this out. The temporal coverage proportion of FY-3B LST cannot achieve 100% temporal coverage at a given location due to the discontinuous revisit orbit. The previous "100%" label was an error in color bar scaling. We have corrected Fig 1. As shown in the revised figure, the actual maximum temporal coverage proportion in our dataset is 92.8%. Furthermore, to avoid any ambiguity, we have explicitly added the definition of "temporal coverage proportion" in Page 2, Lines 82-85.

Fig. 1

"

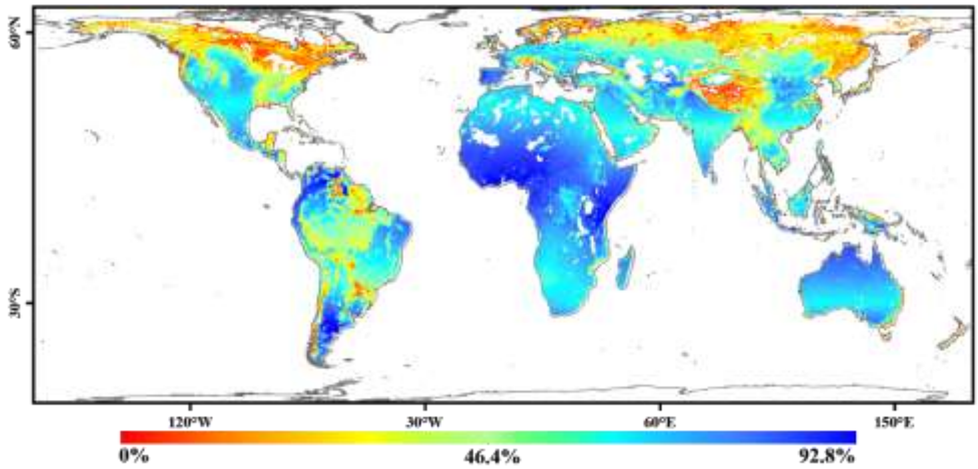

Fig. 1. The temporal coverage proportion of FY-3B SM at the global scale from 12 July 2011 to 19 August 2019."

Page 2, Lines 82-85

"

Fig. 1 shows the temporal coverage proportion of FY-3B SM at the global scale from 12 July 2011 to 19 August 2019. The temporal coverage proportion is defined as the ratio of the number of days with valid satellite retrievals to the total number of days within the study period for each pixel. Due to the discontinuous revisit orbits, the maximum temporal coverage proportion reaches 92.8%. It is seen that the proportion varies spatially, with particularly severe gaps observed in high latitude areas.

"

2) Table 1: The distinction between sparse and dense networks should be clarified. The definition seems to be based on the number of stations within each FY-3B pixel. If this is the case, it would be useful to report the number of matched FY-3B pixels and the average number of stations per pixel, rather than the total number of stations per network.

Re: Thank you for this valuable suggestion. We have clarified the definition of sparse and dense in-situ networks in Page 11, Lines 373-376. Specifically, the classification is based on the spatial distribution of the stations relative to the FY-3B pixels, a network is defined as dense when all stations are located within a single FY-3B pixel, whereas it is defined as sparse when the stations are distributed across multiple FY-3B pixels. We have also added two columns to Table 1, namely, the number of matched FY-3B SM pixels and the average number of stations per pixel. The former indicates the number of FY-3B pixels containing at least one matched in-situ station. The latter is calculated by dividing the total number of stations in each network by the number of matched FY-3B pixels and is rounded to the nearest integer. These additional statistics provide a clearer quantitative description of the spatial concentration of the stations within each network. The corresponding description has been revised in Page 11, Lines 378-381 and Table 1 has been updated accordingly.

Page 11, Lines 373-376

"Meanwhile, to minimize the effects of differences in spatial locations of stations, all networks were divided into two types according to the distribution of stations, including sparse (i.e., stations are distributed across multiple FY-3B SM pixels) and dense (i.e., all stations are distributed within a single FY-3B SM pixel) in-situ data."

Page 11, Lines 378-381

"As shown in Table 1, a larger number of matched FY-3B pixels indicates a more spatially dispersed station distribution, whereas a larger average number of stations per pixel indicates a greater concentration of stations within individual FY-3B pixels."

Table 1

"

Types	***Network***	***Number of matched FY-3B SM pixels***	***Average number of stations per pixel***	***Location (country)***	***Period (yyyy/mm/dd)***
Sparse	***AMMA-CATCH***	*3*	*2*	*Benin, Niger, Mali*	*2006/01/01 to 2018/12/31*
	ARM	*20*	*1*	*USA*	*1993/06/29 to 2025/12/14*
	CTP_SMTMN	*11*	*5*	*China*	*2010/08/01 to 2016/09/19*
	MAQU	*9*	*3*	*China*	*2008/05/13 to 2019/06/01*
	NAQU	*3*	*4*	*China*	*2010/06/15 to 2019/09/12*
	ORACLE	*3*	*1*	*France*	*1985/10/18 to 2013/09/09*
	OZNET	*10*	*2*	*Australia*	*2001/09/12 to 2021/09/01*
	PBO_H2O	*123*	*1*	*USA*	*2004/09/27 to 2017/12/16*
	REMEDHUS	*4*	*5*	*Spain*	*2005/03/15 to 2025/01/01*
	SCAN	*188*	*1*	*USA*	*1996/08/15 to 2025/12/16*
	SMN-SDR	*16*	*2*	*China*	*2018/07/25 to 2019/12/31*
	SNOTEL	*270*	*1*	*USA*	*1996/09/10 to 2025/12/16*
	SOILSCAPE	*5*	*21*	*USA*	*2011/08/03 to 2025/12/16*
	TAHMO	*2*	*1*	*Côte d'Ivoire, Nigeria, Ghana, Uganda, Rwanda, Kenya*	*2015/06/17 to 2022/07/07*
	TxSON	*5*	*8*	*USA*	*2014/10/01 to 2022/11/08*
	USCAN	*93*	*1*	*USA*	*2000/11/15 to 2025/12/16*
	XMS-CAT	*4*	*2*	*Spain*	*2016/08/01 to 2025/12/16*
Dense	***BIEBRZA_S-1***	*1*	*30*	*Poland*	*2015/04/23 to 2018/12/01*
	Ru_CFR	*1*	*2*	*Russia*	*2015/05/25 to 2024/12/31*
	SD_DEM	*1*	*1*	*Sudan*	*2002/02/08 to 2020/11/12*
	SWEX_POLAND	*1*	*6*	*Poland*	*2000/01/01 to 2013/05/06*
	VAS	*1*	*3*	*Spain*	*2010/01/01 to 2012/01/01*

"

3) The manuscript contains too many figures and tables with overlapping messages. Figs. 5–6 and

Tables 2–3 all present in-situ validation. Figs. 7–10 and Table 4 also convey the same simulated gap evaluation and do not provide different messages. Fig. 12 is merely a sequence of images. Many of these materials could be summarized into a figure/table or moved to the supplement.

Re: Thank you for this valuable suggestion. We agree that some of the figures and tables in the original manuscript conveyed related information, and we have therefore carefully reconsidered their presentation to reduce redundancy and improve the readability of the manuscript. Regarding Figs. 5–6 and Tables 2–3, although all four present in-situ validation results, they correspond to two different types of station networks. Specifically, Fig. 5 and Table 2 present the results for sparse networks, whereas Fig. 6 and Table 3 present the results for dense networks. This separation was designed to minimize the effects of differences in spatial locations of stations and to evaluate the reconstruction under different station-distribution conditions. In addition, the dense-network evaluation includes a comparison with the original FY-3B SM, which provides an additional reference for assessing whether the reconstructed dataset can preserve the characteristics of the original product. Therefore, these two sets of results provide complementary information rather than simply repeating the same evaluation. For Figs. 7–10 and Table 4, the four figures correspond to simulated missing-data rates of 30%–40%, 40%–50%, 50%–60%, and 60%–70%, respectively. These figures are primarily intended to provide a visual and qualitative comparison of the reconstructed spatial results by different methods under different missing-data conditions. In contrast, Table 4 provides a quantitative and comprehensive evaluation of the reconstruction accuracy using statistical metrics across the different missing-data scenarios. Thus, the figures and table provide complementary qualitative and quantitative assessments. Following the reviewer's suggestion, we have reduced the amount of material presented in the main text by moving the results for the 30%–40% and 40%–50% missing-data scenarios (Figs. 7 and 8) to the Supplementary Material. Regarding Fig. 12, the 30-day continuous results were selected and presented together to demonstrate the ability of the reconstruction method to preserve the temporal and spatial variations of soil moisture. To reduce the number of figures in the main text, Fig. 12 has also been moved to the Supplementary Material in the revised manuscript.

4) The visual comparisons between soil moisture products are not always convincing. For example, several patch-level figures show similar soil moisture patterns, so it is difficult to judge whether GSP is truly better. Difference maps would be more informative than absolute maps alone.

Re: Thank you for the valuable suggestion. We have comprehensively updated Fig. 14 to include the absolute difference maps for both models (i.e., the "Absolute TTP–Noah difference" and "Absolute GSP–Noah difference" columns). As shown in Fig. 14, the absolute GSP-Noah difference maps exhibit noticeably lower discrepancies compared to the TTP model across the selected regions. To reflect this important enhancement, we have revised Fig. 14 and added the following detailed discussion to Page 25, Lines 566-573 to explicitly analyze the difference maps.

Fig. 14

“

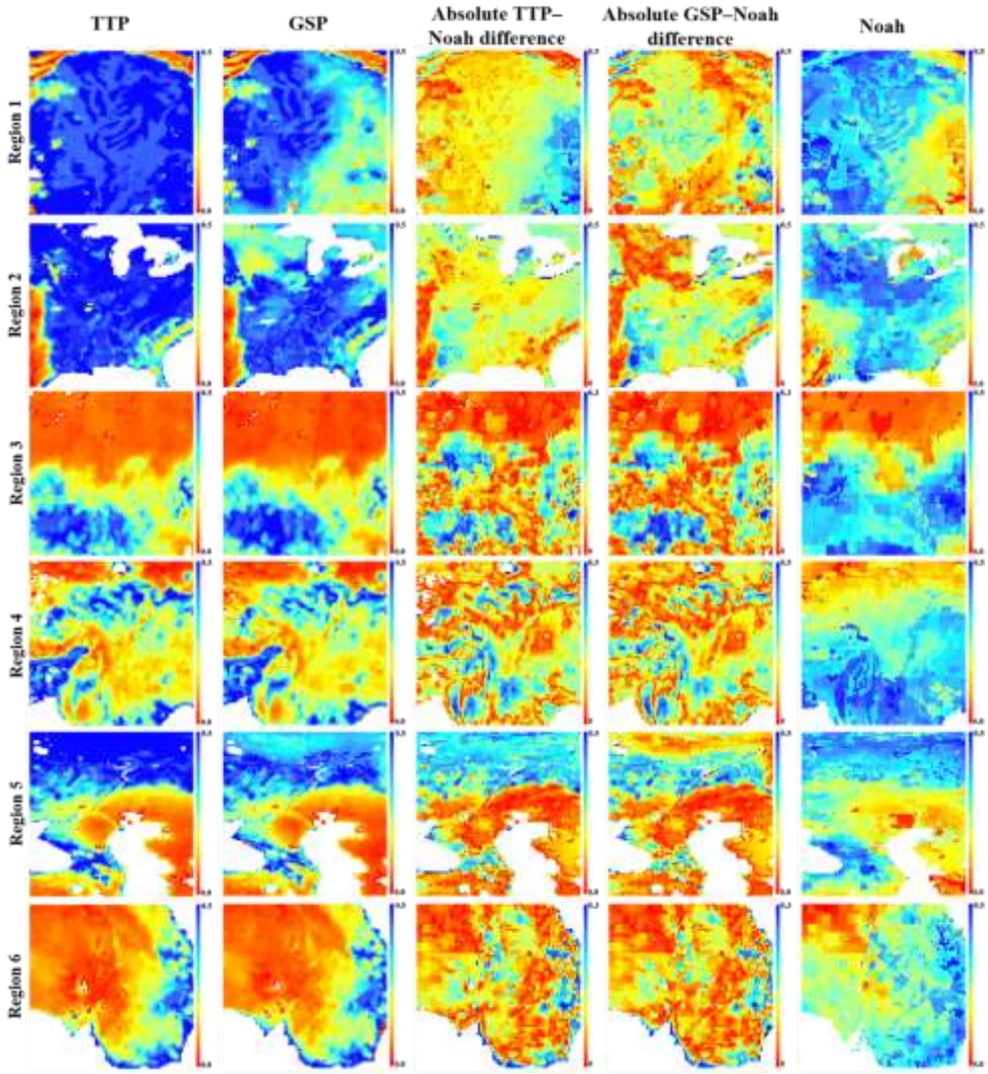

Fig. 14. *Visual comparisons of reconstructed SM spatial patterns and their absolute differences for six randomly selected regions. Columns from left to right show the TTP- and GSP-reconstructed FY-3B SM datasets, the absolute differences of TTP and GSP compared to the GLDAS-Noah reference (Absolute TTP–Noah difference and Absolute GSP–Noah difference), and the GLDAS-Noah SM dataset itself. The latitude and longitude of the patch center as: Region 1 (4.0054° S, 67.4187° W); Region 2 (37.7069° N, 88.2429° W); Region 3 (11.6971° N, 26.2907° E); Region 4 (28.3807° N, 99.1757° E); Region 5 (48.4103° N, 47.1150° E); and Region 6 (28.6028° S, 140.8243° E).*"

Page 25, Lines 566-573

"*To clearly illustrate the model improvements, absolute difference maps against the GLDAS-Noah reference are included. Generally, the TTP-based FY-3B SM reproduces less spatial texture and exhibits larger discrepancies in the difference maps. Conversely, GSP-based results present closer spatial structure and lower absolute error than TTP-based FY-3B SM, emphasizing the importance of considering spatial information in the FY-3B SM reconstruction. Furthermore, as shown in Table 6, the GSP outperforms TTP in terms of all accuracy metrics, with R being 0.0317 larger and ubRMSE being 0.0366 m^3/m^3 lower.*"

5) Please carefully check types. For example, in lines 261-263, “form” should be “from”. In line 383, “FY-3M” should be “FY-3B”. At line 497, 1 February 2011 is outside the study period.

Re: Thank you for carefully pointing out these typographical and consistency issues. We have thoroughly checked the manuscript and corrected all the indicated errors. We have also carefully proofread the manuscript to identify and correct similar typographical errors and inconsistencies throughout the revised manuscript. The corresponding corrections can be found on Page 8, Lines 269–273, Page 12, Lines 405–406, and Page 22, Lines 520–522 of the revised manuscript.

Page 8, Lines 269–273

“It should be noted that, for the special case of periods without sufficient earlier (or later) FY-3B SM data, that is, from 12 July 2011 to 15 July 2011 (16 August 2019 to 19 August 2019), the FY-3B SM data from T to $T+4$ (from $T-4$ to T) are used as input for the model (i.e., treat the data from $T-4$ to $T-1$ (from $T+1$ to $T+4$) as completely missing data).
”

Page 12, Lines 405–406

“In the validation based on the in-situ data at the sparse networks, the 17 sparse networks were used to evaluate the accuracy of the seamless FY-3B dataset generated by different methods.”

Page 22, Lines 520–522

“Fig. 11 shows the original and reconstructed FY-3B SM on five randomly selected dates (i.e., 1 October 2011, 1 August 2013, 1 June 2015, 1 April 2017, and 1 February 2019).”