Response to Referee #3 Comments

We are grateful to the reviewer for the helpful feedback. By taking these suggestions into account, we have revised the MS.

Referee #3 Comments:

The manuscript describes an approach of constructing the global MSS model, with witch the authors have added a new product (SDUST2020) to the market. Also, they have evaluated its quality by comparing it with two popular models (DTU and CLS). I personally support it's publication in ESSD after properly addressing the questions raised by previous reviewers.

I do not have any additional major concerns after reading the comments from the other two reviewers. However, one problem is that I am not sure how the authors are going to improve their manuscript based on those comments. It seems to me that the authors are focusing too much on clarifying their method to just the reviewer him/herself, and did not say anything what has been done to prevent a similar question being raised by a common reader. I therefore suggest the authors update their manuscript lively after responding the reviewers if its possible. If update the manuscript constantly is not an option, at least, they should describe in details how the comments are addressed in the manuscript itself.

After reading the authors reply to the comments I find that the authors have response fairly well. The authors response the reviewers comments fairly well but If a question has been raised or a misunderstanding has been made by an reviewer (expert) when reading the manuscript, then their is good chance that the same will happen to a other readers. Therefore, in their reply, the authors should focus more on describing how they plan to improve their manuscript rather than response to the reviewer him/herself.

Response: Once again, we are particularly grateful for the reviewers careful reading and constructive comments. Thanks very much for your time.

According to the comments and suggestions from Review #1 and #2, we have tried our best to improve the previous manuscript ESSD-2022-178. We have marked all the revisions with track changes on the revised manuscript.

Comments from Referee #1: Whether the altimeter data were retracked? If so, what retracking method was used? And how coastal altimeter data were treated in this study?

Response: Thanks. All the altimetry data used in this study are selected from the along-track Level-2p (L2P; version_02_00) products. They have not been retracked, but they have been preprocessed, including quality control and editing of data to select valid ocean data. The purpose of data preprocessing is to select valid measurements over the ocean with the data editing criteria. The editing criteria are defined as minimum and maximum thresholds for altimeter, radiometer and geophysical parameters (detailed in the along-track L2P product handbook). After data preprocessing, data near the coastline with poor quality have been eliminated (CNES, 2020).

We have added relevant data descriptions to the revised manuscript, please refer to Lines 71-76 in the revised manuscript.

CNES: Along-track level-2+ (L2P) SLA product handbook. SALPMU-P-EA-23150-CLS, Issue 2.0, https://www.aviso.altimetry.fr/fileadmin/documents/data/tools/hdbk_L2P_all_missions_except_S3.pdf, 2020.

Comments from Referee #1: What's the meaning of f(t) in equation (4)? It is suggested not to use the same character for different quantities in equation (3)-(5).

Response: Thanks. f(t) is the systematic errors, which include the radial orbit error, residual ocean variation, residual geophysical corrections, and so on. The same character for different quantities in equation (3)-(5) have been addressed, please refer to Lines 165, and 175-176 for details in the revised manuscript.

Comments from Referee #1: In section 3.1, T/P series data between 66°S and 66°N were used to calculate ocean variability correction for ERS/GM, HY-2A/GM, SARAL and Cryosat-2 which latitude ranges beyond 66°. It need to extrapolate. How does the polynomial fitting interpolation (PFI) perform to do the extrapolation?

Response: Thanks. Since the GM data does not have the characteristics of repeated periods like ERM data, so the ocean variability correction of GM data cannot be addressed by the method of collinear adjustment. Currently, the main methods for the correction of GM data for ocean variability are the objective analysis or based on the use of polynomial functions (e.g. polynomial fitting interpolation, PFI). This study combines these two methods for the ocean variability correction of GM data. The objective analysis method is adopted for the GM data between 66°S and 66°N, while the PFI method is adopted for GM data beyond 66°S or 66°N. In PFI method, seasonal variations are extracted using grid sea level variation time series, interpolated to the GM observations and corrected. The seasonal variations are extracted from the monthly averaged grid sea level variation time series between 1993 and 2019 provided by AVISO, with spatial resolution of $15' \times 15'$.

We have marked the revision with track changes on the revised manuscript. Please refer to Lines 159-167 for details in the revised manuscript.

Comments from Referee #1: In Figure 6-8, there are large differences in polar regions between MSS models. What's the reason?

Response: Thanks. The difference between MSS models depends on the data set used for calculation and the data processing method (Schaeffer et al., 2012). From Figure 6-8, the differences between the three models in the long wavelength are mainly concentrated in the polar regions and the western boundary current region (including the Kuroshio Current, Mexico Gulf, Agulhas Current, etc.). There are two reasons: on the one hand, it is related to the large sea level change in these regions (Jin et al., 2016); on the other hand, it is also related to the different altimeter data used and data processing methods implemented in the modeling (Andersen and Knudsen, 2009; Schaeffer et al., 2012; Pujol et al., 2018). A significant fraction of the large-scale MSS model differences observed in polar regions was shown to originate in different ocean variability corrections or altimeter cross-calibration methods in different MSS models

(Pujol et al., 2018).

We have marked the revision with track changes on the revised manuscript. Please refer to Lines 305-309 for details in the revised manuscript.

Andersen, O. B., and Knudsen, P.: DNSC08 mean sea surface and mean dynamic topography models, J. Geophys. Res.-Oceans, 114(C11), 327-343, https://doi.org/10.1029/2008JC005179, 2009.

Jin, T., Li, J., Jiang, W: The global mean sea surface model WHU2013, Geod. Geodyn., 7, 202-209, http://dx.doi.org/10.1016/j.geog.2016.04.006, 2016.

Pujol, M.-I., Schaeffer, P., Faugère, Y., Raynal, M., Dibarboure, G., and Picot, N.: Gauging the improvement of recent mean sea surface models: a new approach for identifying and quantifying their errors, J. Geophys. Res.-Oceans, 123(8), 5889-5911, https://doi.org/10.1029/2017JC013503, 2018.

Schaeffer, P., Faugére, Y., Legeais, J. F., Ollivier, A., Guinle, T., and Picot, N.: The CNES_CLS11 global mean sea surface computed from 16 Years of satellite altimeter data, Mar. Geod., 35, 3-19, https://doi.org/10.1080/01490419.2012.718231, 2012.

Comments from Referee #2: Altimeter processing description: how were the altimeter data processed and potentially retracked. Were 1 or 20 hz data used for the derivation?. Which range and geophysical corrections were used?. Where the state-of-the-art tide model FES2014b is used consistently.

Response: Thanks. All the altimetry data used in this study were selected from the along-track Level-2p (L2P; version_02_00) products released by the AVISO. The L2P products are generated by the 1 Hz mono mission along-track altimeter data processing segment for Sentinel-3B, Sentinel-3A, Cryosat-2, SARAL/AltiKa, HaiYang-2A, Jason-3, Jason-2, Jason-1, Geosat Follow On, ERS-1, ERS-2, Envisat, and Topex/Poseidon missions. These altimeter data have not been retracked, but they have been preprocessed, including quality control and editing of data to select valid ocean data. The purpose of data preprocessing is to select valid measurements over the ocean with the data editing criteria. The editing criteria are defined as minimum and maximum thresholds for altimeter, radiometer and geophysical parameters (detailed in the along-track L2P product handbook). After data preprocessing, data near the coastline with poor quality have been eliminated (CNES, 2020). Also, all altimetric measurements have been corrected for instrumental errors, environmental perturbations (wet tropospheric, dry tropospheric and ionospheric effects), the ocean sea state bias, the tide effect (ocean tide, solid earth tide and pole tide) and atmospheric pressure (combining atmospheric correction: high frequency fluctuations of the sea surface topography and inverted barometer height correction). The detail of these corrections applied is given in the along-track L2P product handbook (CNES, 2020). The effects of ocean tide for all the altimeter missions are corrected by the ocean tide model of FES2014B.

We have marked the revision with track changes on the revised manuscript. Please refer to Lines 62-76 for details in the revised manuscript.

CNES: Along-track level-2+ (L2P) SLA product handbook. SALPMU-P-EA-23150-CLS, Issue 2.0, https://www.aviso.altimetry.fr/fileadmin/documents/data/tools/hdbk_L2P_all_missions_except_S3.pdf, 2020.

Comments from Referee #2: The crossover adjustment is also severely problematic. When the author writes that the crossover adjustment is carried out in two steps using 1) a condition adjustment method and 2) filtering and prediction along the track. I do think that the second step is a crossover adjustment and that the sentence is largely garbled. Also, the central quantity f is not defined in eq 4.

Response: Thanks. The crossover adjustment is based on the difference between two observations at the same point to integrate different satellite altimeter data (including ERM data and GM data) or to determine corrections to measurements (Huang et al., 2008). The classical crossover adjustment regards the radial orbit error as one of the dominant sources of errors affecting altimeter data and that error can be sufficiently modelled by either a time- or a distance-dependent polynomial (Wagner, 1985; Rummel, 1993). However, the process of solving the equation system in classical crossover adjustment is complicated by a rank deficiency problem, and the computational procedure is complex and cumbersome (Huang et al., 2008). Moreover, in this study, the radial orbit error was of a magnitude similar to that of other physical and geometric uncertainties, such as the inconsistency of the satellite orbit frame and the additional error due to the residual ocean variation and geophysical corrections. To address these limitations, Huang et al (2008) modified the classical crossover adjustment method by dividing it into two steps: (i) condition adjustment at crossover adjustment, and, (ii) filtering and predicting of the observational corrections along each track.

(i) Condition adjustment at crossover adjustment

As we know, the altimeter observation h can be split up into a track-independent part h_0, only depending on the measurement location, and a residual part Δh, which is track dependent. The residual part Δh consists of a part due to the systematic error δh and the stochastic measurement inaccuracy Δ. That is:

$$h = h_0 + \Delta h + \Delta \tag{1}.$$

By introducing the difference of two sea surface height observations at the crossover point of ascending track i and descending track j as the crossover observation, we can define an error equation as:

$$v_{ij}^{a} - v_{ij}^{d} = h_{ij}^{a} - h_{ij}^{d} = d_{ij} \tag{2},$$

where $h_{ij}^{a} - h_{ij}^{d} = d_{ij}$ is the discrepancy at the crossover point $p(i, j)$; the right superscript a indicates ascending tracks, and the right superscript d indicates descending tracks. As for a survey network constructed by M ascending tracks and N descending tracks, the error equations can be expressed in matrix notation as follows:

$$\boldsymbol{BV} - \boldsymbol{D} = \boldsymbol{0} \tag{3},$$

where $\boldsymbol{V}$ represents the correction vector including the signal (systematic error) and noise (random error) parts; $\boldsymbol{B}$ is the coefficient matrix which consists of 1 and -1; $\boldsymbol{D}$ indicates the discrepancy vector. The least square solution of Eq. (3) is:

$$\boldsymbol{V} = \boldsymbol{P}^{-1}\boldsymbol{B}^{\mathrm{T}}\left(\boldsymbol{BP}^{-1}\boldsymbol{B}^{\mathrm{T}}\right)^{-1}\boldsymbol{D} \tag{4}.$$

The cofactor matrix is:

$$\boldsymbol{Q}_{V} = \boldsymbol{P}^{-1}\boldsymbol{B}^{\mathrm{T}}\left(\boldsymbol{BP}^{-1}\boldsymbol{B}^{\mathrm{T}}\right)^{-1}\boldsymbol{BP}^{-1} \tag{5},$$

where $\boldsymbol{P}$ is the weighting matrix of discrepancy observations. Suppose the sea surface height observations to be independent along each track, Eq. (4) can be further rewritten as:

$$v_{ij}^a = p_{ij}^d d_{ij} / \left(p_{ij}^a + p_{ij}^d \right) \quad (6),$$

$$v_{ij}^d = -p_{ij}^a d_{ij} / \left(p_{ij}^a + p_{ij}^d \right) \quad (7),$$

where p_{ij}^a and v_{ij}^a represent the weight factor of observation and its correction along ascending track i at crossover point $p(i, j)$, respectively; p_{ij}^d and v_{ij}^d represent the weighting factors of observation and its correction along descending track j at crossover point $p(i, j)$, respectively.

(ii) Filtering and predicting along tracks

According to the modern adjustment theory, after the observational correction vector is calculated from Eq. (4), it can be further considered as a new kind of observations and then be filtered using an error model. Taking into account the fact that the amplitude of orbit error is, now, almost the same as that of influence of other physical and geometric uncertainties such as the inconsistency in the satellite orbit frame, and the additional errors caused by residual ocean variation and various physical corrections, a reasonable error model is constructed to illustrate the change of signals, with which the filtering and prediction of crossover adjustment corrections are done along each single track. It is clear from the error analysis that the performance of the errors from satellite altimetry appears mainly to have a systematic influence on measurements. The combined effect of the errors will vary in very complicated ways. It may consist of linear, periodic, and irregular trends. After finishing a series of tests using general polynomial and trigonometric polynomial error models, it is found that a combined model of general and trigonometric polynomials is more advantageous in describing change of systematic errors in satellite altimetry. This model can be expressed as follows:

$$f(t) = a_0 + a_1 \cdot (t - T_0) + \sum\nolimits_{(j=1)}^{n} (b_i \cdot \cos(j \cdot \omega \cdot (t - T_0)) + c_i \cdot \sin(j \cdot \omega \cdot (t - T_0))) \quad (8)$$

where $f(t)$ is the systematic errors; t is the observation time of the sea surface height; a_0, a_1, b_i, and $c_i (i = 1, \cdots, n)$ are model parameters to be solved; ω represents the angular frequency corresponding to the duration of a surveying track ($\omega = 2\pi/(T_1 - T_0)$, where T_0 and T_1 represent the start and end times of the surveying track, respectively); and n is a positive integer determined by the length of the track. Based on empirical evidence, n is proposed to be 1–2 for a short track, 3–5 for a middle-long track, and 6–8 for a long track (Huang et al., 2008).

After condition adjustment at crossover points, a new error equation can be constructed with error model (8) at each crossover point as follows:

$$v = f(t) + \Delta \quad (9).$$

And its matrix form is:

$$\boldsymbol{V} = \boldsymbol{AX} + \boldsymbol{U} \quad (10),$$

where $\boldsymbol{V}$ is the virtual observation vector; $\boldsymbol{U}$ is the correction vector of the virtual observations; $\boldsymbol{A}$ is a known coefficient matrix and is expressed as

$$\boldsymbol{A} = \begin{bmatrix} 1 & t - T_0 & \cos\omega(t - T_0) & \sin\omega(t - T_0) & \cdots & \cos(m\,\omega(t - T_0)) & \sin(m\,\omega(t - T_0)) \end{bmatrix} \quad (11);$$

$\boldsymbol{X}$ is the vector of the undetermined coefficient and is expressed as

$$\boldsymbol{X} = \begin{bmatrix} a_0 & a_1 & c_1 & b_1 & \cdots & c_m & b_m \end{bmatrix}^{\mathrm{T}} \quad (12).$$

The least squares solution of Eq. (10) is

$$\hat{\boldsymbol{X}} = (\boldsymbol{A}^T \boldsymbol{P}_V \boldsymbol{A})^{-1} \boldsymbol{A}^T \boldsymbol{P}_V \boldsymbol{V} \quad (13),$$

where $\boldsymbol{P}_V$ is the weight matrix of virtual observations.

The estimated parameter vector $\hat{X}$ is put into Eq. (8). According to the observation time of the along-track sea surface height of the track, the residuals of sea surface height systematic errors can be calculated and corrected by Eq. (8).

The crossover adjustment method used in this study is not a new methodology. It has been described in detail by Huang et al. (2008) and Yuan et al. (2020). We have marked the revision with track changes on the revised manuscript. Please refer to Lines 171-172 for details in the revised manuscript.

Huang, M., Guan, Z., Zhai, G., and Ouyang, Y.: On the compensation of systematic errors in marine gravity measurements, Marine Geodesy, 22(3), 183-194. http://dx.doi.org/10.1080/014904199273452, 1999.

Huang, M., Zhai, G., Ouyang, Y., Lu, X., Liu, C., and Wang, R.: Integrated data processing for multi-satellite missions and recovery of marine gravity field, Terr. Atmos. Ocean. Sci., 19, 103-109, https://doi.org/10.3319/TAO.2008.19.1-2.103(SA), 2008.

Rummel, R.: Principle of satellite altimetry and elimination of radial orbit errors. In: Rummel, R., Sansò, F. (Eds.), Satellite Altimetry In Geodesy And Oceanography. Springer, Berlin, Heidelberg, Germany, pp.190-241, https://doi.org/10.1007/BFb0117929, 1993.

Wagner, C.A.: Radial variations of a satellite orbit due to gravitational errors: implications for satellite altimetry, Journal of Geophysical Research: Solid Earth 90(B4), 3027-3036, https://doi.org/10.1029/JB090iB04p03027, 1985.

Yuan, J., Guo, J., Liu, X., Zhu, C., Niu, Y., Li, Z., Ji, B., and Ouyang, Y.: Mean sea surface model over China seas and its adjacent ocean established with the 19-year moving average method from multi-satellite altimeter data, Cont. Shelf Res., 192(1), 104009, https://doi.org/10.1016/j.csr.2019.104009, 2020.

Comments from Referee #2: interesting enough I noticed large discrepancies with other MSS models above 80N. This could fit with the fact that 80N is the northern limit of the 20x20 degree boxes, so data in the few degrees to the north of 80N are not adjusted?

Response: Thanks. The difference between MSS models depends on the data set used for calculation and the data processing method (Schaeffer et al., 2012). As show in Figures 6-8, the differences between the three models in the long wavelength are mainly concentrated in the polar regions and the western boundary current region (including the Kuroshio Current, Mexico Gulf, Agulhas Current, etc.). There are two reasons: on the one hand, it is related to the large sea level change in these regions (Jin et al., 2016); on the other hand, it is also related to the different altimeter data used and data processing methods implemented in the modelling (Andersen and Knudsen, 2009; Schaeffer et al., 2012; Pujol et al., 2018). A significant fraction of the large-scale MSS model differences observed in polar regions is shown to originate in different ocean variability corrections or altimeter cross-calibration methods in different MSS models (Pujol et al., 2018).

We have marked the revision with track changes on the revised manuscript. Please refer to Lines 305-309 for details in the revised manuscript.

Andersen, O. B., and Knudsen, P.: DNSC08 mean sea surface and mean dynamic topography models, J. Geophys. Res.-Oceans, 114(C11), 327-343, https://doi.org/10.1029/2008JC005179, 2009.

Jin, T., Li, J., Jiang, W: The global mean sea surface model WHU2013, Geod. Geodyn., 7, 202-209, http://dx.doi.org/10.1016/j.geog.2016.04.006, 2016.

Pujol, M.-I., Schaeffer, P., Faugère, Y., Raynal, M., Dibarboure, G., and Picot, N.: Gauging the improvement of recent mean sea surface models: a new approach for identifying and quantifying their errors, J. Geophys. Res.-Oceans, 123(8), 5889-5911, https://doi.org/10.1029/2017JC013503, 2018.

Schaeffer, P., Faugére, Y., Legeais, J. F., Ollivier, A., Guinle, T., and Picot, N.: The CNES_CLS11 global mean sea surface computed from 16 Years of satellite altimeter data, Mar. Geod., 35, 3-19, https://doi.org/10.1080/01490419.2012.718231, 2012.

Comments from Referee #2: Section 5.1 present the comparison with CLS15 and DTU18 models. Here the authors present the central table 5 which is used to infer the accuracy of the models from high to low. In my view, it only explains that the authors are doing something wrong in my view. First of all the DTU15 and CLS18 MSS are not different on average by 1.27 cm Many investigations (e.g., Pujol et al. 2019) show much smaller numbers. The differences in Table 5 between the model's present standard deviation of >29 centimeters are clearly not what other authors present.

Response: Thanks. The results listed in Table 5 are the statistical results of the comparison between these three models in global ocean. A total of 1 5533 0402 grid points are counted, including grid points in the coastal regions. After outliers in the difference are rejected by three times STD to avoid contamination by the poor observations around coastal regions. The results (shown in Table 4) are consistent with other authors.

Table 4 Statistical results of comparisons between different mean sea surface models after rejecting outlies in the differences by three times STD (Unit: m)

Model discrepancy	Max	Min	Mean	STD	RMS	Number of points
SDUST2020-CLS15	0.0413	-0.0396	0.0009	0.0135	0.0135	133495409
SDUST2020-DTU18	0.0554	-0.0405	0.0074	0.0160	0.0176	131613306
CLS15-DTU18	0.0487	-0.0365	0.0060	0.0142	0.0155	129765806

We have marked the revision with track changes on the revised manuscript. Please refer to Lines 290-297 for details in the revised manuscript.

Comments from Referee #2: Also, the following figure 6 demonstrates that the standard deviation is far less than 29 cm. I think that the authors have potentially forgotten to apply the confidence mask in the Arctic Ocean and elsewhere.

Response: Thanks. Figures 6-8 show the differences in sea surface height between these three models in long and short wavelengths. Long and short wavelengths are selected similar to Andersen et al. (2018) at a wavelength of 150 km as the dividing line. It can be seen from the Figures 6, 7, and 8 that there are no significant differences between these models in the short wavelength (wavelength less than 150 km), and the average differences are within 2 cm, while there are some significant differences in the long wavelength (wavelength greater than 150 km). The differences between these models in the long wavelength are mainly concentrated in the polar regions and the western boundary current region (including the Kuroshio Current, Mexico Gulf, Agulhas Current, etc.). There are two reasons: on the one hand, it is related to the large sea level change in these regions (Jin et al., 2016); on the other hand, it is also related to the different altimeter data used and data processing methods implemented in the modelling (Andersen and Knudsen, 2009; Schaeffer et al., 2012; Pujol et al., 2018). A significant fraction of the large-scale MSS model differences observed in polar regions is shown to originate in different ocean variability corrections or altimeter cross-calibration methods in different MSS models (Pujol et al., 2018).

We have marked the revision with track changes on the revised manuscript. Please refer to Lines 305-309 for details in the revised manuscript.

Andersen, O. B., and Knudsen, P.: DNSC08 mean sea surface and mean dynamic topography models, J. Geophys. Res.-Oceans, 114(C11), 327-343, https://doi.org/10.1029/2008JC005179, 2009.

Andersen, O. B., Knudsen, P., and Stenseng, L.: A new DTU18 MSS mean sea surface– improvement from SAR altimetry, In: 25 Years of Progress in Radar Altimetry Symposium, Portugal, 2018.

Jin, T., Li, J., Jiang, W: The global mean sea surface model WHU2013, Geod. Geodyn., 7, 202-209, http://dx.doi.org/10.1016/j.geog.2016.04.006, 2016.

Pujol, M.-I., Schaeffer, P., Faugère, Y., Raynal, M., Dibarboure, G., and Picot, N.: Gauging the improvement of recent mean sea surface models: a new approach for identifying and quantifying their errors, J. Geophys. Res.-Oceans, 123(8), 5889-5911, https://doi.org/10.1029/2017JC013503, 2018.

Schaeffer, P., Faugére, Y., Legeais, J. F., Ollivier, A., Guinle, T., and Picot, N.: The CNES_CLS11 global mean sea surface computed from 16 Years of satellite altimeter data, Mar. Geod., 35, 3-19, https://doi.org/10.1080/01490419.2012.718231, 2012.

Comments from Referee #2: A little later the authors also present the average and RMS about the formal error (again garbled sentence) of 1 and 1.5 cm for SDSUT. What does this mean and how does it relates to Table 5.

Response: Thanks. We apologize for the poor language of our manuscript. The formal error is caused by the three terms: an instrumental noise, a residual effect of the oceanic variability, and an along-track bias, and obtained at the optimal interpolation output. These three terms are complementary and correspond, respectively, to a white noise, a spatially correlated noise (at mesoscale wavelengths), and a long-wavelength error that is assumed to be constant along the tracks. The formal error does not match the precision of the MSS but is nonetheless an excellent indicator of the consistency of the grid (Schaeffer et al., 2012; Pujol et al., 2018).

In practice, this formal error variance corresponds to a local minimum in the least squares sense; it depends on the spatial distribution and the density of the data used in the suboptimal estimation, but also on the noise budget. Overall, the map of this formal error gives us information about the homogeneity of the solution, and more locally the ratio between grid points is close to the relative accuracy (Pujol et al., 2018).

We have marked the revision with track changes on the revised manuscript. Please refer to Lines 316-322 for details in the revised manuscript.

Pujol, M.-I., Schaeffer, P., Faugère, Y., Raynal, M., Dibarboure, G., and Picot, N.: Gauging the improvement of recent mean sea surface models: a new approach for identifying and quantifying their errors, J. Geophys. Res.-Oceans, 123(8), 5889-5911, https://doi.org/10.1029/2017JC013503, 2018.

Schaeffer, P., Faugére, Y., Legeais, J. F., Ollivier, A., Guinle, T., and Picot, N.: The CNES_CLS11 global mean sea surface computed from 16 Years of satellite altimeter data, Mar. Geod., 35, 3-19, https://doi.org/10.1080/01490419.2012.718231, 2012.

Comments from Referee #2: The comparison with tide gauges is questionable. First of all. Have the author included the formal error on the MSS in this comparison and does the MSS fit within this?. Secondly, have the authors ensured that the same version of the reference ellipsoid (TOPEX vs WGS84/GRS80) has been used and that the version is employing the tide system?. In this section the authors only present numbers but no interpretation of the results. Is it realistic that the differences range up to nearly a meter (with a formal error of 1 cm claimed for the SDSUT.

Response: Thanks. As mentioned in the answer to the previous question, the formal error does not match the precision of the MSS but is nonetheless an excellent indicator of the consistency of the grid. Here, we compare the sea surface heights of these three models (CLS15, DTU18, and SDUST2020) with those obtained by GPS-levelled tide gauges around Japan, respectively, to independently validate the accuracy differences of these models in coastal regions. In this comparison, the sea surface heights obtained by GPS-levelled tide gauges has been adjusted to have the same reference ellipsoid as T/P. In table 6, the STD of sea surface heights difference between MSS model and the GPS-levelled tide gauges reaches decimeter level. The reason is may be closely related to the poor observations of offshore altimeter data.

We have marked the revision with track changes on the revised manuscript. Please refer to Lines 334-346 for details in the revised manuscript.