Response to referee #1

Title: A global long-term (1981–2019) daily land surface radiation budget product from AVHRR satellite data using a residual convolutional neural network

MS_No: ESSD-2021-250

Thanks very much for taking your time to review this manuscript. We really appreciate all your valuable comments and constructive suggestions! The specific responses to your all comments are listed below one by one.

General comments:

Comment 1:

Getting finer resolution from coarse resolution data is not easily accepted without clear explanation. You explain that "spatial adjacent effect" is accomplished by applying reanalysis data and angular information of the satellite measurement and solar position. The spatial adjacent effect seems to be the novel advantage of the CNN method over other existing methods, but the explanation lacks detail and examples to help the reader better understand the upscaling process. Does it use reanalysis vertical profiles to correct the path of upwelling radiation to the satellite? Or is it some kind of statistical approach?

Response 1:

We use CNN model to upscale the *in-situ* measurements at "points" to a 0.05° spatial resolution. We selected these sites whose measurements can well represent the average state of surface R_n at a 5-km geospatial extent using the ETC method. Errors of the upscaling process can be weakened to a certain degree because of a good spatial representation of selected ground-based measurements within AVHRR footprint. MERRA2 reanalysis has a spatial resolution of 0.5°×0.625°. Therefore, MERRA2 data are resampled to the 0.05° resolution using the nearest neighbor method to avoid introducing new errors.

thresholds between 0.2 and 0.9 at intervals of 0.1, a threshold of 0.9 was selected, above which the sites were assigned as 'reliable' (also see Sect. 5). Errors of upscaling reliable site-based measurement can be weakened to a certain degree due to its better representativeness within the AVHRR footprint.

We do not use the reanalysis vertical profiles to correct the path of upwelling radiation to the satellite. Instead, we provide comprehensive information within a determined optimal geospatial extent for CNN to automatically extract the most important features related to reliable site-based R_n measurements based on multiple filters. The influence of spatial adjacent effect on surface radiation is highly related to the viewing geometry of sun-target-sensor (Wang et al., 2017), surface, and atmospheric conditions, e.g., the presence of clouds (Wyser et al., 2002), by multiple scattering, reflection, and absorption in the entire atmosphere column on pixel scale. Surface net radiation is generally inferred from satellite-observed radiance based on the independent pixel approximation in the past retrieval algorithms. However, with the increased spatial resolution, the spatial adjacent

effects (or 3-D radiative effects) caused by clouds, water vapor, and aerosols become more significant and are not ignored in the inversion process. To address the spatial adjacent effect on surface R_n, a proper geospatial extent centered on the site was determined by the MLR method using AVHRR TOA observations. Comprehensive surface and atmospheric information within the determined spatial extent is necessarily considered in the inversion process of surface R_n. CNN is a tool to properly process the input data in a form of the multi-dimensional matrix. Therefore, CNN can extract the most essential feature from the input spatial data within the determined spatial extent to relate with site-based measurements, which is better than using the information on individual pixels. The operation weakens spatial adjacent effect on the surface R_n at the center pixel to a certain degree. The input features include AVHRR TOA observations representing comprehensive surface and atmospheric information, viewing geometry, and MERRA2 R_n addressing the difference in temporal scales. In this way, more accurate R_n values are obtained at the center pixel. More explanations are included in the revised manuscript.

As the spatial resolution of satellite sensors increases, the spatial adjacent effects induced by spatially inhomogeneous atmospheric constitutes (or clouds) fields become more significant, for example, clouds affect the distribution of surface radiation in a region larger than the resolution of an individual pixel. One spatial adjacent effect is the diffusion of radiation that removes part of radiation from an atmospheric column and transfer it to neighboring columns. Two other effects are related to the solar and viewing geometry, such as a shift of the apparent position of clouds and their shadows. Surface R_n is no longer accurately estimated with retrieval algorithms based on the individual pixel approximation (IPA). Comprehensive information within a certain spatial extent centered at reliable sites needs to be applied to help retrieve surface R_n. CNN model can extract features hierarchically from input multi-channel images using multiple filters. Therefore, the most important feature information regarding reliable site-based R_n measurements can be effectively extracted by CNN within a certain spatial extent rather than on IPA, to help retrieve R_n, which weakens the spatial adjacent effects to a certain extent.

Specific comments:

Comment 1:

14 Include the temporal resolution of the Rn estimates here.

Response 1:

Thanks for your suggestion. The temporal resolution of the Rn estimates has been included in line 14 as follows.

network (RCNN) integrating spatially adjacent information to improve the accuracy of retrievals. A global high-resolution (0.05°), long-term (1981–2019), and daily mean R_n product was subsequently generated from Advanced Very High-Resolution Radiometer (AVHRR) data. Specifically, the RCNN was employed to establish a nonlinear relationship between globally

---

Comment 2:

20-24 The statement beginning with “Inter-comparisons with three…” is not true. In section 4.3.3 you state: “The validation results in Fig. 8 and Table 7 for the ice/snow surface cover type further confirm that GLASS Rn product may offer a better performance in Greenland region.”

Response 2:

We are extremely grateful for you to point out the problem. We have revised the statement in the revised manuscript as follows.

and bias of 0.84, 26.66 Wm^{-2} (31.66%), and 1.59 Wm^{-2} (1.89%), respectively. Inter-comparisons with three other R_n products, i.e., the 5 km Global Land Surface Satellite (GLASS), the 1° Clouds and the Earth's Radiant Energy System (CERES), and the 0.5° × 0.625° Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2), illustrate that our AVHRR R_n retrievals have the best accuracy under most of the considered surface and atmospheric conditions, especially thick cloud or hazy conditions. However, the performance of the model needs to be further improved for the snow/ice cover surface. The spatiotemporal analyses of these four R_n datasets indicate that the AVHRR R_n product reasonably replicates the

---

Comment 3:

40 “radiation” is not needed in front of “radiometers”

Response 3:

Thanks for your nice suggestion. The “radiation” has been deleted in the revised manuscript.

ground-based measurements are widely used to study spatiotemporal variations in regional surface radiation and to evaluate gridded products (Jia et al., 2018; Zhang et al., 2020; Zhang et al., 2015). Nevertheless, the high cost of maintaining ~~radiation~~ radiometers means that stations are sparely distributed, severely hindering our ability to study and understand the

---

Comment 4:

132-161 This section describes the instruments used in the various networks which range from good thermopile pyranometers and pyrgeometers to not-so-good net radiometers. You only provide performance measures for the thermopile pyranometers, which are generally good. You don't provide any performance information on the net radiometers, which are notoriously bad, especially the REBS model. According to Table 2, net radiometers dominate your observational dataset. You should provide performance measures of the net radiometers. Also, Table 2 is incomplete. For some you specify “Eppley PIR,” and others just “Eppley.”

Response 4:

Thanks for your kind suggestion. After extensively reading literature and corresponding websites, the uncertainty of each instrument is listed as follows, and the corresponding content is also included in the revised manuscript.

To be specific, the operational thermoelectric pyranometers are known for their high-accuracy

performance, with a spectral response of 0.3-3.0 μm, a sensitivity of 7-14μVW^{-1} m^2, a thermal effect of less than 5%, and an annual stability of 5% (Lu et al., 2011; Jiang et al., 2019). The Eppley Precision Infrared Radiometers (PIR, 3.5-50 μm) and Kipp & Zonen CG 4 pyrgeometers (4.5-42 μm) are applied to measure the surface radiation with a uncertainty of ± 6% or 15 Wm^{-2} at the 95% confidence level (Philipona et al., 1998). The largest uncertainty for surface radiation measurements is ~2% for pyrheliometers and ~5% for pyranometers (i.e., 15 Wm^{-2}), respectively (Augustine et al., 2000). Additionally, the radiation measurements obtained by Kipp & Zonen CNR1 and CNR4 instruments are with an expected accuracy of ±10% for daily totals (Wang and Dickinson, 2013). The radiation observations measured by Kipp & Zonen net radiometers (CNR1, 5-50 μm or CNR1-lite, 4.5-42 μm), are with uncertainty of ~10% at 95% confidence level for daily totals (Yamamoto et al., 2005). Besides, the uncertainties of the shortwave radiation measured by LI-COR Photodiode and R_n observed by REBS Q*7 are about 5 (5-15%) and 10 Wm^{-2} (5-50%), respectively, at monthly time scale (Box and Rinke, 2003; Steffen and Box, 2001).

To be specific, the operational thermoelectric pyranometers are known for their high-accuracy performance, with a spectral response of 0.3-3.0 μm, a sensitivity of 7-14μVW^{-1} m^2, a thermal effect of less than 5%, and an annual stability of 5% (Lu et al., 2011; Jiang et al., 2019b). The Eppley Precision Infrared Radiometers (PIR, 3.5-50 μm) and Kipp & Zonen CG 4 pyrgeometers (4.5-42 μm) are applied to measure the surface radiation with a uncertainty of ± 6% or 15 Wm^{-2} at the 95% confidence level (Philipona et al., 1998). The largest uncertainty for surface radiation measurements are ~2% for pyrheliometers and ~5% for pyranometers (i.e., 15 Wm^{-2}), respectively (Augustine et al., 2000). Additionally, the radiation measurements obtained by Kipp & Zonen CNR1 and CNR4 instruments are with an expected accuracy of ±10% for daily totals (Wang and Dickinson, 2013). The radiation observations measured by Kipp & Zonen net radiometers (CNR1, 5-50 μm or CNR1-lite, 4.5-42 μm), are with uncertainty of ~10% at 95% confidence level for daily totals (Yamamoto et al., 2005). Besides, the uncertainties of the shortwave radiation measured by LI-COR Photodiode and R_n observed by REBS Q*7 are about 5 (5-15%) and 10 Wm^{-2} (5-50%), respectively, at monthly time scale (Box and Rinke, 2003; Steffen and Box, 2001). To deal with equipment and operational errors, daily mean surface R_n measurements were calculated based on several strict

Network/Program	Instrument	Temporal Interval	Number of sites
ARM	Kipp&Zonen CNR-1	10 minutes	34
AsiaFlux	Kipp&Zonen CNR-1/EKO MS201	30 minutes	31
BSRN	Kipp&Zonen CG4/Eppley. PIR	1 minutes	21
CEOP	Eppley. PIR/EKO MS202	30 minutes	16
CEOP-Int	Kipp&Zonen CG4/Eppley. PIR	30 minutes	8
ChinaFlux	Kipp&Zonen CNR-1	30/60 minutes	3
EOL	Kipp&Zonen pyrgeometers, Eppley. PIR	30/60 minutes	17
GCNET	Li Cor Photodiode & REBS Q*7	60 minutes	18
GAME.ANN	EKO MS0202F	30 minutes	3
Global FluxNet	Kipp&Zonen CNR-1, etc.	30 minutes	314
HiWATER	Kipp&Zonen CNR-1/CNR-4	10 minutes	19
IMAU-Ktransect	Kipp&Zonen CNR-1	60 minutes	4
LBA-ECO	Kipp&Zonen CG2/CNR-1	30 minutes	8
PROMICE	Kipp&Zonen CNR-1/CNR-4	10 minutes	24
SURFRAD	Eppley pyrgeometer	1/3 minutes	7

Comments 5:

By “thermal effect,” are you referring to the thermal offset of single black detector pyranometers? If so, there are references for this measurement error.

Response 5:

The “thermal effect” refers to the thermal offset of the thermopile pyranometers. The measurements error is referred to studies of Jiang et al. (2019) and Lu et al. (2011)

Comment 6:

169. What does “along with inverse navigation to relate a specific Earth location to each sensor’s instantaneous field of view” mean?

Response 6:

The sentence of “along with inverse navigation to relate a specific Earth location to each sensor’s instantaneous field of view” means geometric correction which is one of the three components of the AVHRR Land Pathfinder II processing system. The other two components are radiometric in-flight vicarious calibrations for the visible and near-infrared channels and atmospheric correction, respectively (Pedelty et al., 2007). Specifically, navigation is a process that relates an Earth location to an instantaneous field of view (IFOV) of the sensor. The inverse navigation refers that the nearest IFOV scan number and position are determined for each grid cell of a predetermined geographic grid, which is a preprocessing for generating a consistent, long-term AVHRR data set at a resolution of 0.05° (El Saleous et al., 2000).

Comment 7:

180-194 In this description of the GLASS product, Rn is estimated from downward shortwave radiation, and other variables using multiple MARS learners. Where do the input data come from?

Response 7:

The shortwave radiation, albedo, and NDVI data are from GLASS products (Liang et al., 2020; Xiao et al., 2017). Other meteorological varibales come from MERRA2 reanalysis (Gelaro et al., 2017). The related information is included in the revised manuscript.

index (NDVI). Multiple MARS learners were employed to establish efficient statistical relationships using GLASS downward shortwave radiation and MERRA2 meteorological variables, allowing land surface R_n to be estimated from these inputs across most spatial domains (Jiang et al., 2016; Jiang et al., 2015). Conversely, when surface solar radiation data were not available,

Comment 8:

145 217 The last phrase of this sentence "the diurnal variation of daily surface Rn." Does not make sense.

Response 8:

Thanks for your nice suggestion. The last phrase of "the diurnal variation of daily surface Rn" has been deleted in the revised manuscript.

to other reanalysis data. Therefore, MERRA2 R_n data calculated from four surface radiative components were also used in this study to help retrieve accurate high-resolution surface R_n estimates by providing average atmospheric information ~~about the diurnal variation of daily surface R_n.~~

150

Comment 9:

285 What does "when deeper networks converge" mean?

Response 9:

155 A deeper network converages, meaning the training and test errors no longer decrease with increasing training epochs. However, a degradation problem may expose that training accuray gets saturated and then degrades rapidly with network depth increasing. In other words, adding more layers to a suitably network leads to higher training error (He and Sun, 2015; Srivastava et al., 2015), though the deeper network starts converaging. The residual learning framework proposed by He et al. (2016) is thus applied
160 to deal with the degradation problem.

Comment 10:

323 This sentence does not make sense. Do you mean "Reliable and unreliable sites from each
165 observation network, separated by a threshold ETC-derived correlation coefficient of 0.9, are listed in Table 5"?

Response 10:

Thanks for your kind suggestion. Your understanding of this sentence is right! We have revised the sentence according to your expression. The number of reliable and unreliable sites for each observation
170 network, identified by a threshold of 0.9 for the ETC-derived correlation coefficient, is listed in Table 5.

~~The reliable sites and unreliable sites identified a threshold of 0.9 for the ETC-derived correlation coefficient are summarized in Table 5 for each observation network.~~ The number of reliable and unreliable sites for each observation network, identified by a threshold of 0.9 for the ETC-derived correlation coefficient, is listed in Table 5. A total of 275 sites could be considered reliable, accounting for ~48% of the sites. Furthermore, no site was considered reliable for some observation

Comment 11:

334 - 337 Please include references for the ARM, SURFRAD, BSRN, and FluxNet networks.

Response 11:

We are grateful for the suggestion. The corresponding references for the four networks are included in the revised manuscript. ARM (Stokes and Schwartz, 1994), SURFRAD (Augustine et al., 2000), BSRN (Ohmura et al., 1998), and FluxNet (Wilson et al., 2002).

networks. In contrast, some of the international observational networks, such as BSRN (Ohmura et al., 1998) and FluxNet (Wilson et al., 2002), provide many ground-based measurements with sufficient spatial representativeness for R_n at 5 km resolution. In addition, the ARM (Stokes and Schwartz, 1994) and SURFRAD (Augustine et al., 2000) networks were classified as containing reliable sites. In situ measurements from the SURFRAD (Augustine et al., 2000) network were well

Comment 12:

I assume the color bar represents a normalized count scaled to the most frequent count. Regardless, explain the color bar in the caption.

Response 12:

Thanks for your suggestion. The color bar illustrates the normalized density of samples. The corresponding explain has been included in the caption of the figure.

Figure 5: Scatterplots of (a) mode training (fitting) accuracy and (b) model test accuracy for the reliable training and independent validation sites. The color bar illustrates the normalized density of samples.

Comment 13:

414 “under snow and ice surfaces” ? Perhaps use “for snow and ice surfaces” ?

Response 13:

Thanks for your suggestion. We have revised the phrase according to your comment.

PROMICE network for most of the large sites in the GCNET and PROMICE networks identified as unreliable sites. Thus, the RCNN model has less knowledge of R_n dynamics ~~under~~ for snow and ice surfaces. The most significant difference for RMSE was observed over the ARM network, for which the mean RMSE value decreased by 2.1 Wm^{-2} for the AVHRR R_n retrievals

Comment 14:

453 Change phrasing to “…especially clouds that have significant impacts on shortwave…”

Response 14:

Thanks for your suggestion. We have revised the phrase according to your comment.

clouds and CWV control surface R_n dynamics under cloud-sky conditions, ~~especially clouds, which have significant impacts on surface shortwave and longwave radiation~~ especially clouds that have significant impacts on shortwave and longwave radiation. Therefore, AOD, CWV, and cloud optical thickness (COT, as a surrogate for cloud optical properties) were

Comment 15:

454 Where do cloud optical thickness (COT) and cloud water vapor (CWV) data come from?

Response 15:

The COT and CWV data come from MERRA2 reanalysis. We have added the information in the revised manuscript.

radiation. Therefore, AOD, CWV, and cloud optical thickness (COT, as a surrogate for cloud optical properties) derived from MERRA2 were employed to analyze the sensitivity of the accuracy of the AVHRR and GLASS R_n retrievals to variations in these influencing factors. In addition, R_n retrieval performance at different elevations was also evaluated.

Comment 16:

462 The sentence beginning with "Therefore, the performance…" does not make sense. Perhaps the end of that sentence should read: "…is comparable with regard to the accuracy of their Rn retrievals."

Response 16:

Thanks for your kind suggestion. We have revised the phrase according to your comment.

clear-sky conditions, which results in surface total solar radiation dominated direct solar radiation. Therefore, the performance of the RCNN model and the MARS models used for the GLASS R_n product (Jiang et al., 2016) is comparable ~~the accuracy of the R_n retrievals~~ with regard to the accuracy of their R_n retrievals. However, when the absorption and scattering effects are enhanced for direct solar radiation from TOA, depending on the individual pixel approximation (IPA), it is difficult to retrieve

Comment 17:

512 – 519 I don't understand your Figure 11. The AVHRR and GLASS Rn's as a function of COT are nearly on top of each other, yet the bias plotted on the same charts is significant. What am I missing here? Regardless, in the caption please define the bias and what the shading represents.

Response 17:

The absoulate bias is defined as the absoluate difference between daily mean AVHRR and GLASS R_n, i.e., $\left|Rn_{avhrr} - Rn_{glass}\right|$. To more easily understand, the bias is replaced by the difference. The shading represents the variation range (stand deviation) of global daily AVHRR and GLASS R_n retrievals and their absoulate differences. The related information has been added in the figure caption.

Figure 11: Variations in the spatial and temporal consistency of AVHRR and GLASS daily R_n retrievals against cloud optical thickness (COT) in (a) January and (b) July 2008. The absolute difference is defined as $|Rn_{avhrr} - Rn_{glass}|$. The shading represents the variation range (stand deviation) of global daily AVHRR and GLASS R_n retrievals and their absolute differences.

The distributions of the absolute differences between daily mean AVHRR and GLASS R_n values in January and July 2008 are shown in Figure 1. There exist large absolute differences in both January and July. It is reasonable that large absolute differences occur in Figure 11 in the manuscript. The reason for the AVHRR and GLASS R_n values being nearly on top of each other (solid lines) is the effect of averaging operation over multiple land pixels within a certain COT range. However, the large variation range (shading) of daily R_n retrievals means a single pixel within a certain COT range may have a large absolute difference under some specific conditions.

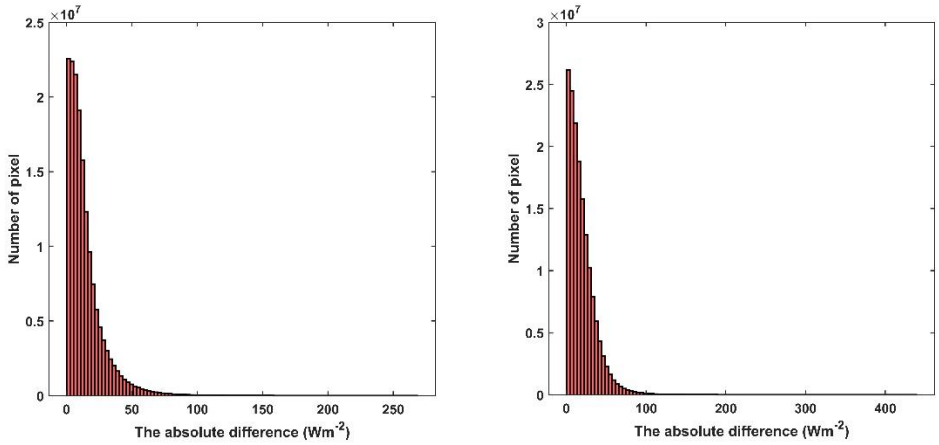

Figure 1: The distributions of absolute differences in January (left) and July (right), respectively.

Comment 18:

524 Please state how the difference is defined. Jan.- July or July – Jan.

Response 18:

The difference of cloud fraction is defined as Jan.-July. We have included the information in the revised manuscript.

information includes the properties of the entire cloud layer. Figure S2 shows the spatial distribution of the monthly mean cloud cover fraction (CF) at the global scale in January and July, and the corresponding differences in CFs (Jan.-July). In January, the CFs are higher over most land regions except in northern Africa, southern America, northern Austria and southern

Comment 19:

525 Do you mean “northern Australia” and “South America” ?

Response 19:

Note the areas where the CFs in July is larger than that in January 2008 (blue areas, Figure 2). The areas include Central Africa, Southern Asia, Southern Australia and Antarctica. We have revised the corresponding part in the manuscript.

cloud cover fraction (CF) at the global scale in January and July, and the corresponding differences in CFs (Jan.-July). In January, the CFs are higher than in July over most land regions except in ~~northern Africa, southern America, northern Austria and southern Asia~~ Central Africa, Southern Asia, Southern Australia and Antarctica. However, most regions had smaller CFs

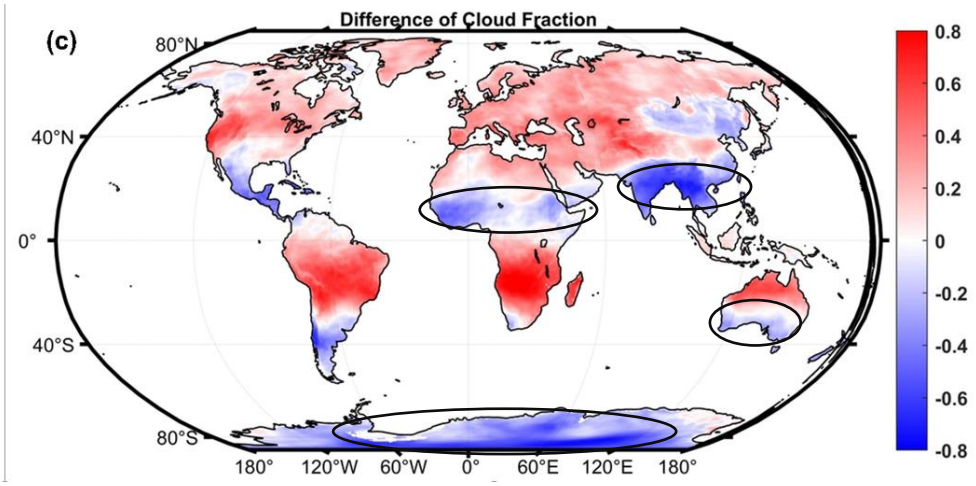

Figure 2: Spatial distribution of the global CFs differences between in January and in July 2008.

Comment 20:

545 "produced by NOAA" ? Should this read "replaced by NOAA"?

Response 20:

Thanks for your careful work. We have revised the mistake in the revised manuscript.

periods correspond to the alternative update times of the NOAA-series satellites. For example, NOAA-11 was successfully ~~produced~~ succeeded by NOAA-14 from 1994 to 1995. Similarly, NOAA-16 replaced NOAA-14 in 2000 for monitoring of the Earth's surface and atmosphere. During these periods of satellite replacement, the corresponding AVHRR data contain large

Comment 21:

555-560 The four timeseries in Fig. 12 after 2017 for all data sets may be well correlated but are obviously wrong and could not be used for climate studies. What does the shading represent in Fig. 12?

Response 21:

The LTDR project only uses afternoon satellite to generate the long-term AVHRR dataset because the atmospheric correction algorithm would produce high uncertainty when applied to low sun elevation pixels from the morning (am) satellites. Afternoon satellites include NOAA7, NOAA9, NOAA11, NOAA14, NOAA16, NOAA18 and NOAA19 (Figure 3). The use of these satellites alone inevitably

leads to small gaps in the data in exchange for a higher accuracy in the atmospheric correction. The time series is not fully complete and presents some observational gaps. The most important two were found in 1994 and from 2018 onwards. In the first case, important gaps and noise were found in the images from March to September and empty data from September to December, due to NOAA11 orbital degradation. From 2018 onwards the data quality has been degrading due to important gaps in the images and the presence of artefacts (Otón et al., 2021). This is why the R_n timeseries after 2017 seem to be abnormal. At several studies, authors suggest that 1994, 2018, 2019, and 2020 are not used due to the poor quality of AVHRR data (Hansen et al., 2020; Tian et al., 2015). The corrsponding contents were included in the revised manuscript.

The shading represents the variation range of global monthly mean R_n. The information has been included in the revised manuscript.

Figure 12: Long-term temporal variation of (a) monthly average R_n and (b) monthly R_n anomalies for the AVHRR, CERES, GALSS and MERRA2 datasets, respectively. The shading represents the variation range (stand deviation) of the global monthly mean R_n.

Note that the LTDR project only use afternoon satellite to generate the AVHRR product to do with the high uncertainty of the atmospheric correction algorithm when applied to low sun elevation pixels present in morning (am) satellites. Afternoon satellites include NOAA-7, NOAA-9, NOAA-11, NOAA-14, NOAA-16, NOAA-18, NOAA-19, and NOAA-20. The use of these satellites alone inevitably leads to small gaps in the data in exchange for a higher accuracy in the atmospheric correction. The time series is not fully complete and presents some observational gaps. Specifically, some large discrepancies occur,

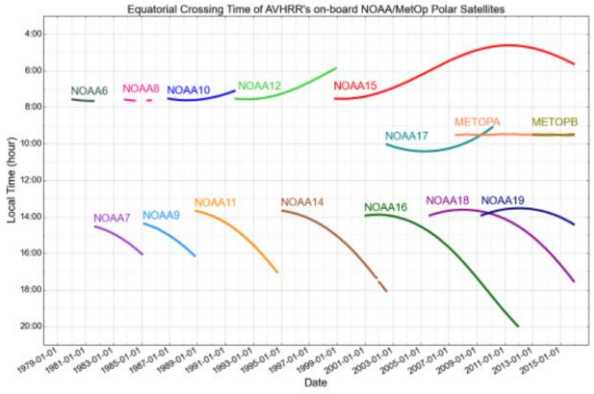

Figure 3: Local overpass time of all NOA satellites containing the AVHRR sensor. Figure obtained from Clerbaux et al. (2020).

Comment 22:

557 Has the AVHRR calibration across all satellites been applied to the AVHRR data shown in Fig. 12?

Response 22:

The LTDR product performs geolocation, calibration, and atmospheric and surface anisotropy correction for all AVHRR sensors aboard the NOAA afternoon (pm) satellites (Vermote and Saleous, 2006; Vermote and Kaufman, 1995; Otón et al., 2021; Franch et al., 2017). The calibration method proposed by Vermote and Kaufman (1995) is applied consistently across the AVHRR instruments onboard various NOAA satellites. Relevant information refers to

https://landweb.modaps.eosdis.nasa.gov/cgi-bin/ltdr/ltdr/ltdrPage.cgi?fileName=avhrr_calib.

Comment 23:

572 Do you mean 0.708?

Response 23:

Thanks for your carefule examination. I looked back at the statistics, and the best R is 0.708, not 7.08. I have revised the mistake in the manuscript.

correspond to approximately 15 × 15 km^2 (B3) to 135 × 135 km^2 (B19) on the ground. The results are shown in Fig. 13. Overall, the average R increases from 0.61 to ~~7.08~~ 0.708, and RMSE decreases from 50.12 to 46.17, respectively, for the MLR model. As the valid spatial extent increases, essential and complete spatial features are exposed and incorporated into the MLR model,

Comment 24:

591 What is a wide overpass time?

Response 24:

The wide overpass time refers to a broad range of local time for satellite crossing as shown in Fig. 3 from 13:00 to 20:00 LT for afternoon satellites.

Comment 25:

600 What Study?

Response 25:

Shupe et al. (2011) found annual cloud occurrence fractions are 58%–83% at the Arctic observatories, with a clear annual cycle wherein clouds are least frequent in the winter and most frequent in the late summer and autumn. We have included the study in the revised manuscript.

about the diurnal cycles of the atmosphere and clouds is more important for daily surface radiation estimation at high latitudes than that at middle and low latitudes. Shupe et al. (2011) found annual cloud occurrence fractions are 58%–83% at the Arctic observatories, with a clear annual cycle wherein clouds are least frequent in the winter and most frequent in the late summer and autumn.

Technical corrections:

Comment 26:

164 You don't need ", respectively" in this sentence.

Response 26:

Thanks for your suggestion. I have deleted the ", respectively" in the revised manuscript.

AVHRR TOA observations at five spectral channels (a visible band (0.55–0.68 μm), a near-infrared band (0.75–1.1 μm), a middle-infrared band (3.55–3.93 μm), and two thermal bands (10.5–11.3 and 11.5–12.5 μm~~, respectively~~) were utilized for

335

Comment 27:

"BSRN_DRA" site.

340 **Response 27:**

Thanks for your suggestion. I have revised the "BSRA_DRA" to "BSRN_DRA".

MERRA2 and CERES-SYN R_n retrievals show higher values compared to the in situ measurements at the ~~BSRA_DRA~~ BSRN_DRA site, especially during 140–200 day period. In comparison, the AVHRR and GLASS R_n values closely match the

345 **Comment 28:**

483 "very low" not "vary low"

Response 28:

Thanks for your suggestion. We have revised the "vary low" to "very low".

1 Wm^{-2}. The lower accuracy of AVHRR R_n values for these two elevation ranges is attributable to the less reliable sites used for the RCNN training. In addition, the AVHRR R_n retrievals show steady and ~~vary~~ very low (close to zero) biases under different conditions, while the biases of the GLASS R_n retrievals show a high degree of variation. This illustrates that the

350

Comment 29:

543 1999-2000

Response 29:

355 Thanks for your careful examination. We have corrected the mistake.

some large discrepancies occur, during some periods including 1994–1995, 1999–2000, 2007–2008, and 2018–2019. These periods correspond to the alternative update times of the NOAA-series satellites. For example, NOAA-11 was successfully

References

360 *Augustine, J. A., DeLuisi, J. J., and Long, C. N.: SURFRAD–A national surface radiation budget network*

for atmospheric research, Bulletin of the American Meteorological Society, 81, 2341-2358, 2000.

Box, J. E. and Rinke, A.: Evaluation of Greenland ice sheet surface climate in the HIRHAM regional climate model using automatic weather station data, Journal of Climate, 16, 1302-1319, 2003.

Clerbaux, N., Akkermans, T., Baudrez, E., Velazquez Blazquez, A., Moutier, W., Moreels, J., and Aebi, C.: The Climate Monitoring SAF Outgoing Longwave Radiation from AVHRR, Remote Sensing, 12, 929, 2020.

El Saleous, N., Vermote, E., Justice, C., Townshend, J., Tucker, C., and Goward, S.: Improvements in the global biospheric record from the Advanced Very High Resolution Radiometer (AVHRR), International Journal of Remote Sensing, 21, 1251-1277, 2000.

Franch, B., Vermote, E. F., Roger, J.-C., Murphy, E., Becker-Reshef, I., Justice, C., Claverie, M., Nagol, J., Csiszar, I., Meyer, D., Baret, F., Masuoka, E., Wolfe, R., and Devadiga, S.: A 30+ Year AVHRR Land Surface Reflectance Climate Data Record and Its Application to Wheat Yield Monitoring, Remote Sensing, 9, 296, 2017.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., and Reichle, R.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), Journal of climate, 30, 5419-5454, 2017.

Hansen, M., Song, X., DiMiceli, C., Carroll, M., Sohlberg, R., Kim, D., and Townshend, J.: MEaSURES Vegetation Continuous Fields ESDR Algorithm Theoretical Basis Document (ATBD) Version 2.0, 2020.

He, K. and Sun, J.: Convolutional neural networks at constrained time cost, Proceedings of the IEEE conference on computer vision and pattern recognition, 5353-5360,

He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778, 10.1109/CVPR.2016.90, 2016.

Jiang, H., Lu, N., Qin, J., Tang, W., and Yao, L.: A deep learning algorithm to estimate hourly global solar radiation from geostationary satellite data, Renewable & Sustainable Energy Reviews, 114, 109327, 2019.

Liang, S., Cheng, J., Jia, K., Jiang, B., Liu, Q., Xiao, Z., Yao, Y., Yuan, W., Zhang, X., and Zhao, X.: The Global LAnd Surface Satellite (GLASS) product suite, Bulletin of the American Meteorological Society, 1-37, 2020.

Lu, N., Qin, J., Yang, K., and Sun, J.: A simple and efficient algorithm to estimate daily global solar radiation from geostationary satellite data, Energy, 36, 3179-3188, 2011.

Ohmura, A., Dutton, E. G., Forgan, B., Fröhlich, C., Gilgen, H., Hegner, H., Heimo, A., König-Langlo, G., McArthur, B., and Müller, G.: Baseline Surface Radiation Network (BSRN/WCRP): New precision radiometry for climate research, Bulletin of the American Meteorological Society, 79, 2115-2136, 1998.

Otón, G., Lizundia-Loiola, J., Pettinari, M. L., and Chuvieco, E.: Development of a consistent global long-term burned area product (1982–2018) based on AVHRR-LTDR data, International Journal of Applied Earth Observation and Geoinformation, 103, 102473, https://doi.org/10.1016/j.jag.2021.102473, 2021.

Pedelty, J., Devadiga, S., Masuoka, E., Brown, M., Pinzon, J., Tucker, C., Vermote, E., Prince, S., Nagol, J., Justice, C., Roy, D., Junchang, J., Schaaf, C., Jicheng, L., Privette, J., and Pinheiro, A.: Generating a Long-term Land Data Record from the AVHRR and MODIS Instruments, 2007 IEEE International Geoscience and Remote Sensing Symposium, 23-28 July 2007, 1021-1025, 10.1109/IGARSS.2007.4422974,

Philipona, R., Fröhlich, C., Dehne, K., DeLuisi, J., Augustine, J., Dutton, E., Nelson, D., Forgan, B.,

Novotny, P., and Hickey, J.: The Baseline Surface Radiation Network pyrgeometer round-robin calibration experiment, Journal of Atmospheric and Oceanic Technology, 15, 687-696, 1998.

Shupe, M. D., Walden, V. P., Eloranta, E., Uttal, T., Campbell, J. R., Starkweather, S. M., and Shiobara, M.: Clouds at Arctic atmospheric observatories. Part I: Occurrence and macrophysical properties, Journal of Applied Meteorology and Climatology, 50, 626-644, 2011.

Srivastava, R. K., Greff, K., and Schmidhuber, J.: Highway networks, arXiv preprint arXiv:1505.00387, 2015.

Steffen, K. and Box, J.: Surface climatology of the Greenland ice sheet: Greenland Climate Network 1995–1999, Journal of Geophysical Research: Atmospheres, 106, 33951-33964, 2001.

Stokes, G. M. and Schwartz, S. E.: The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the cloud and radiation test bed, Bulletin of the American Meteorological Society, 75, 1201-1222, 1994.

Tian, F., Fensholt, R., Verbesselt, J., Grogan, K., Horion, S., and Wang, Y.: Evaluating temporal consistency of long-term global NDVI datasets for trend analysis, Remote Sensing of Environment, 163, 326-340, 2015.

Vermote, E. and Kaufman, Y.: Absolute calibration of AVHRR visible and near-infrared channels using ocean and cloud views, International Journal of Remote Sensing, 16, 2317-2340, 1995.

Vermote, E. and Saleous, N.: Calibration of NOAA16 AVHRR over a desert site using MODIS data, Remote sensing of Environment, 105, 214-220, 2006.

Wang, K. and Dickinson, R. E.: Global atmospheric downward longwave radiation at the surface from ground-based observations, satellite retrievals, and reanalyses, Reviews of Geophysics, 51, 150-185, 2013.

Wang, T., Shi, J., Husi, L., Zhao, T., Ji, D., Xiong, C., and Gao, B.: Effect of solar-cloud-satellite geometry on land surface shortwave radiation derived from remotely sensed data, Remote Sensing, 9, 690, 2017.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., and Field, C.: Energy balance closure at FLUXNET sites, Agricultural and Forest Meteorology, 113, 223-243, 2002.

Wyser, K., O'Hirok, W., Gautier, C., and Jones, C.: Remote sensing of surface solar irradiance with corrections for 3-D cloud effects, Remote Sensing of Environment, 80, 272-284, https://doi.org/10.1016/S0034-4257(01)00309-1, 2002.

Xiao, Z., Liang, S., Tian, X., Jia, K., Yao, Y., and Jiang, B.: Reconstruction of long-term temporally continuous NDVI and surface reflectance from AVHRR data, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10, 5551-5568, 2017.

Yamamoto, S., Saigusa, N., Gamo, M., Fujinuma, Y., Inoue, G., and Hirano, T.: Findings through the AsiaFlux network and a view toward the future, Journal of Geographical Sciences, 15, 142-148, 2005.