General Comments:

The authors developed 40-year (1980-2019) Taiwan reanalysis downscaling data with a 2-km grid from the ERA5's hourly data using the WRF model. They also validated their results with observations from CWA stations, 1-km TCCIP data and the shortwave radiation data derived from Himawari-8 satellite by CWA. The TReAD data is freely available for climate related studies and would be a very valuable data set for Taiwan. However, the reviewer has comments on procedures, data sources, and analyses used for validation for the authors to consider. Plausible explanations on errors statistics would be very useful.

Specific Comments;

1. Since TCCIP is one of the major data sources for validation. This data set should be properly described in Section 2.3, including variables, time intervals, resolution, accuracies, and periods each variable is available. The only information on this data set is given in Weng and Yang (2018). In the reference list, Weng and Yang (2018) is on daily grided rainfall data during 1960-2015 in Taiwan. Furthermore, this is a gray paper and may not be readily available to readers.
2. In Section 2.1, in addition to model domains, physical parameterization schemes, the authors described the 1-km terrain data and land use data used. There is no information on soil type data used. The soil type is an important parameter for lower boundary conditions over land. Also, there is no mention of SST data used. Did the authors use SST data from ERA5?
3. In Section 2.2, the authors describe the time integration procedures using the WRF model, initialized by the ERA5 data. The authors cited Lynch and Huang (1992) for the digital filter initialization method used. Lynch and Huang (1992) is not cited in alphabetical order in the reference list. A brief description of this initialization technique and blending procedures should be given. In addition to initialization of atmospheric parameters, were any initializations being made for the lower boundary conditions over land?
4. The model has a one-month spin up period before 12-h integrations each 6 hours. For each 12-h integration, the simulation for first 6-h integration is not used to construct the climate time series of downscaled data. What parameters were spin up during the one-month spin up period? Why?
5. Section 2.3: With a 2-km grid, the terrain is supposed to be well represented in the model. What are the differences in the actual terrain height and model terrain height at the 5 mountainous CWA stations?
6. Section 2.3: For shortwave radiation data, for the CWA data derived from satellite for 2011-2018, what is the grid size of this data set? For the shortwave radiation data observed by CWA stations, what are the periods that station observed data are available?
7. Section 3: The order to present validation of ECVs is personal preference. However, since the surface p, T are affected by the terrain height, and surface winds are affected by

the diurnal heating cycle, the reviewer suggests surface p, 2-m T and RH to be presented first followed by surface winds. Localized rainfall distribution and shortwave radiation are affected by locally driven diurnal winds; thus, the reviewer also suggests that these two parameters being discussed after discussion of surface winds.

8. Section 3.3, Validation of surface pressure
   The correlation coefficients for surface pressure are slightly lower for 5 mountain stations (Fig. 13a), the authors suggest that this may be related to elevation differences between the model grids and stations even after the former is adjusted. The elevation differences for these stations are not given. It is not apparent how adjustments to the surface pressure are made. Fig. 13b shows large RMSE values for Yushan (~9 hPa), Alisan (> 2 hPa), Tainan (> 2hPa), Chenggoang (> 6 hPa). These values are large. In addition, a lower coefficient with large RMSE (>3 hPa) were reported at Lanyu; the authors speculate that the model may not catch the exact location of that station due to the fact it is located within a small island off the southeastern Taiwan coast (L331-332). The authors should investigate this possibility based on the data. During the passage of a TC, unlike the other three stations (Fig. 14), the simulated surface pressure at Yushan is persistently lower (more than 12 hPa) than observed; the authors suggested that this may be due to the lower intensity of the TC simulated. This explanation may not be valid as the surface pressure simulated for the three other stations agrees well with observations.
   A few questions for the authors. With a 2-km grid, what are the deficiencies of terrain height for five mountainous stations (Line 166-167) and Lanyu in the model? How would these differences affect the deficiencies in surface pressure at these stations? The deficiencies at Yushan, Chenggong are relatively large with large RMSE, any reasons for the relatively large deficiencies? The TC intensity seems to be well predicted for three stations (Figs. 14a,b,c) but why not at Yushan (Fig. 14d)?
9. Section 3.1, Validation of 2-m T
   Anomalies of long-term trend of mean temperature agree well with CWA station observations (Fig. 6), the long-term mean daily 2-m T also agrees well with TCCIP data (Fig. 8a). However, there are deficiencies. For mean daily 2-m T, there are persistently warm bias over western plain and northern Taiwan, and cold bias over the mountainous areas especially on the eastern slopes as compared with TCCIP 1-km data (Fig. 7). These errors tend to cancel out in the mean temperature. The RMSE errors of 2-m T are relatively large at Yushan (~2 degrees C) even after terrain adjustment and with a lower coefficient (~ 0.7) (Fig. 5). Also, the daily mean maximum 2-m T is persistently 1.5 degrees colder than TCCIP data (Fig. 8b). The reasons for these deficiencies should be discussed. Would the deficiencies affect the simulated daytime upslope flow? Moreover, the descriptions for TCCIP 1-km 2-m T data are not properly given (L177-178).
10. Section 3.6, Validation of RH (%)
    The reviewer presumes the authors validate the 2-m RH (%). The long-term trend of simulated 2-m RH (%) anomalies agrees well with those from the CWA station data (Fig.

24). Thus, there is a mean warming and drying trend over the island which is interesting. From RMSE errors (Fig. 23b) and daily time series plots during the diurnal cycle for a few stations (Fig. 25), it is apparent that the major cities over the western plain and northern Taiwan (Taichung, Chiayi, Tainan, Hsinchu, Taipei) have a relatively large RMSE errors (Fig. 23b) with lower RH (%) especially under the summer monsoon during the daytime (Taichung in Fig. 25). With warmer T and drier RH, how would that affect the downscaled surface winds, precipitation and shortwave radiation over the island?

11. Section 3.5, Validation of wind velocity. The reviewer presume wind velocity is a vector. However, Figs. 19 and 20 only show wind speeds. No information on wind direction is given. CWA station data shows decreasing wind speed over the island but not in the downscaled TReAD data (Fig. 20). No explanation is given. Fig. 21 shows the differences in wind vectors between the 25-km ERA5 data and the downscaled 2-km TReAD data. The comparison is not informative as the ERA-5 data cannot resolve the island induced circulation and winds over the coastal waters. The ERA5 data represents mesoscale winds over Taiwan. Perhaps, the authors may want to compare the downscale 2-km winds in the afternoon vs before sunrise to show the capability of the TReAD data in simulating the island induced circulation over the island interior and coastal waters throughout the year.
12. Section 3.2, Validation of precipitation. Were the hourly rainfall data or daily rainfall accumulation from TReAD and TCCIP data used in this section? TCCIP data is available only during 1960-2015 (Weng and Yang 2018), when compare with TCCIP data, did you use the TReAD data for the same period? Correlation coefficients with CWA rainfall data are low compared to other ECVs (Fig. 9a). The RMSE errors from all stations are > 10 mm which is relatively large as compared with the observed daily rainfall accumulation which may be rather typical (Fig. 9b). The anomalous annual precipitation for TReAD and CWA are consistent (Fig. 10) but with smaller anomalies for TReAD. Since the authors suggest that rainfall over Taiwan is affected by terrain effects, the authors should discuss how do the terrain effects affect temperature, RH, winds over land and coastal waters and precipitation? Are those effects captured by TReAD 2-km data in some way? In Fig. 11, does the figure show the average daily rainfall accumulation for TReAD and TCCIP during the same 1960-2015 period? The figure caption states the bottom row is the difference between TReAD and CWA data. Can you explain the CWA data used? Do you use the daily rainfall from CWA stations? Could you adjust the color table? What are the main deficiencies in daily rainfall accumulation from TReAD? Would that impact the accuracy of the TReAD shortwave radiation data?
13. Section 3.4, Validation of shortwave radiation. The units used in Fig.18 and other figures are not consistent. Any reasons for the use of different units? Are shortwave radiation data related to rainfall in some way? In compassion of anomalies of annual shortwave radiation data between TReAD and CWA station-based data between 1980-2019, there are differences (Fig. 16). The CWA based data used for comparison should be described.

The seasonal means of the shortwave radiation data from CWA, TReAD and ERA5 from 2011-2018 were compared and given in Fig. 17. What is the horizonal spacing for CWA satellite derived data? Information on the data set used should be given. It is apparent that the shortwave radiation data from TReAD is not very useful with significant deficiencies. No explanation is given.

In summary, the reviewers recommend the authors improve descriptions for procedures, data sources, and analyses used for validation. Plausible explanations on errors statistics would also be very useful.