Report on Tasmin and Hurs in ISIMIP3b

Bruno Lemke, Matthias Otto and Tord Kjellstrom

13 June 2021

Executive Summary

In this report we investigate the un-physical condition of Relative Humidity (RH) greater than 100% at minimum temperatures in a large number of 0.5x0.5degree grid cells in ISIMIP3b (46% of the grid cells over the 12 months). Our initial analysis showed that over half these grid cells where RH>100% were at altitudes >400m and/or at temperatures < 0C. We excluded these high altitude and very cold cells from our analysis because of uncertainties with our formulas and data at low temperature and altitude.

Of the remaining datapoints the ISIMIP3b UKesm and the W5E5 data where cell RH at minimum temperature was greater than 100% were very similar suggesting it was not an issue of bias over-correction this time. In comparing W5E5 with weather station data, for grid cells that had weather stations, we could not establish whether minimum temperatures were too low or the dew point temperatures were too high. We also found that RH>100% at minimum temperature for a large number (3343) of weather stations during at least one month between 2001 and 2010.

When we investigated HOURLY weather station data, in NO case did the hourly data have RH>100% at any temperature even for stations where the daily RH at minimum temperature was consistently greater than 100%. This led us to confirm a theory that on the occasions when the drop in temperature at night drags down the dew point so that the daily averaged humidity (average relative humidity, dew point temperature or specific humidity) from hourly humidity can easily lead to the "un-physical" condition where RH might be more than 100% at the daily minimum temperature (which does not fit with physical science).

Introduction

Table 1 Terminology used in this report for ISIMIP3bV2 and ISIMIP3bV3 data*		
Variable	**Our terminology**	**ISIMIP terminology**
Maximum daily temperature (at surface & bias corrected)	Tmax	tasmaxAdjust
Average daily temperature	Tmean	tasAdjust
Minimum daily temperature	Tmin	tasminAdjust
Average daily relative humidity	RH	hurs
Daily specific humidity	SH	huss
Relative humidity at Tmax	RH(Tmax)	
Relative humidity at Tmean	RH(Tmean)	
Relative humidity at Tmin	RH(Tmin)	
Dew Point temperature	Tdew	
*Note: we have run tests on ISIMIP3b historic on three occasions: V1 was before the official version was released (approx. March 2020). This was then replaced by the main version which we call V2 that came out about June 2020. The current version where Tmax and Tmin were corrected we call version V3 which came out in May 2021. Data downloaded from esg.pik-potsdam.de/search/isimip and vre2.dkrz.de:8080/thredds/catalog/catalog.html		

After our identification of extremely high maximum temperatures in the June 2020 version of the ISIMIP3b data set and the successful correction of this by PIK, I was asked by Dr Lange if I could identify the cause of the un-physical condition where RH > 100% at minimum temperature. In Dr Lange's paper where he discussed the bias correction and downscaling method used for ISIMIP (Lange 2019) he talks often about his methods producing data within “the confinement of extreme values to the physically plausible range”, yet an RH>100% is not physically plausible except under super-saturation conditions in an extremely clean atmosphere (Lawrence 2005).

Over a 12 month period using 2001 to 2010 average values we found the RH>100% at minimum temperatures in a large number of 0.5x0.5 degree grid cells in ISIMIP3b for all climate models. We reported this to Dr Lange and he requested us to follow this up. The air temperature at RH = 100% is called the dew point as at that temperature water begins to condense out of the air and dew forms (Lawrence 2005). So Tdew > Tmin equates to RH > 100% according to the calculation formula.

While our heat impact research does not involve Tmin, it does involve RH as heat stress on workers is a combination of RH and Tmax. In the course of our work we have occasionally used Tmin and it was noted that Tmin for many grid cells was less than Tdew. We use Tdew as a humidity measure because this is what weather stations use, while ISIMIP uses RH and SH and UK Climate Research Unit (CRU 2021) uses vapour pressure. ERA5 climate re-analysis data (Copernicus 2021) also uses Tdew. These humidity terms are interchangeable with formulae (Vaisala 2020). We avoid using RH in our impact studies because it is temperature dependent while all the other humidity units listed are temperature independent. In this report we will primarily use RH as it is the one used in ISIMIP3b.

No problems were found with the RH(Tmean) or RH(Tmax) because RH is less than 100% in all grid cells for all months for those two temperatures – presumably because this is corrected to stop RH(Tmean) going over 100% in the bias correction. Note that this was discussed in ISIMIP *Fast Track* where the RH at Tmean was not initially pegged to a maximum of 100%.

One of the questions asked by PIK was whether the reason RH(Tmin)>100% was due to too low Tmin values or too high Tdew values as calculated from RH(Tmean).

To understand why RH(Tmin) > 100% we also investigated data from weather stations in cells where RH(Tmin)>100% and were surprised to find that this unphysical condition also occurs in daily weather station data.

Method

Method 1: 10/30 year monthly averaged data

As ISIMIP3b has 3 main temperatures as output: Tmax, Tmean and Tmin there should therefore be 3 RH measures RH(Tmin), RH(Tmean) and RH(Tmax). To get RH(Tmin) we assume that the average daily RH (as outputted in ISIMIP) occurs at Tmean. To convert RH(Tmean) to Tdew we use the Tetens formula (Lawrence 2005)). The RH results were checked on TWO online calculators – one at NOAA (2021) and the other at Omni Calculator (2021). It should be noted that if the Tdew is greater than the air temperature these calculators generate an error message.

So we obtain Tdew from average daily RH and daily Tmean. Then, knowing Tdew is independent of temperature we can reverse this calculation to find the RH(Tmax) and RH(Tmin). There is one uncertainty in this calculation: we assume Tmean occurs at the same time as daily mean RH. This simultaneous measure is required to calculate Tdew. There is no guarantee that Tmean and RH (as supplied by ISIMIP) occurs at the same time so in this study we used SH (Specific Humidity – which is already independent of temperature). This SH humidity variable is supplied by ISIMIP in the 3b version. When we used ISIMIP3b version of SH to calculate Tdew the results for the number of grid cells where RH(Tmin) > 100% were almost identical to that in deriving Tdew from RH(Tmean) so we used the latter in this report as we already had data for this.

Initially we did NOT use daily data because we wanted to capture the effect over the baseline period and did not concern ourselves with daily fluctuations. We used monthly averages over 30 years (1981-2010) or 10 years (2001-2010). We are aware that this averaging process blunts extremes, but if the 30-year or 10-year average of RH(Tmin) is above 100% then a substantial number of days must have a RH(Tmin) > 100%.

The process for the initial scoping exercise was this: using 30 year averages of Tdew and Tmin, firstly, all Tmin < 1C were removed because we were concerned that the Tdew formula may not apply over ice. We chose <1C rather than <0 C to allow for common small errors in Tmin. To identify the un-physical cells, only grid cells where RH(Tmin) >110% were considered to allow for a 10% error in the RH calculations.

After this scoping exercise more detailed results were obtained by comparing with daily weather station data in grid cells that had weather stations. Daily weather station data were also averaged monthly for the 30 years 1981-2010 and only weather stations with >90% completeness during that time period were used. The weather station data from HothapsSoft (Otto 2020) was, in turn, derived from NOAA daily weather station data: NOAA Global Surface Summary of the Day (NOAA GSOD 2021). When comparing ISIMIP3b data with weather stations, we excluded high altitudes (>400m) and Tmin <1C because of uncertainties with our formulas and data. For example, in mountainous terrain the average grid cell altitude was sometimes thousands of metres different than that from weather stations in the same grid cell. At a Tmin < 0C frost forms rather than dew so we chose an lower limit of 1C for Tmin rather than 0C to cover errors of about 1C in the data.

Method 2: Daily and Hourly weather station data

We used daily (see above) and detailed hourly (or 3 hourly) weather station data to confirm our weather station theory why RH(Tmin)>100%. Daily and hourly data from 5 weather stations in 2020 from the NOAA integrated surface hourly database (NOAA ISH (2021) were used. These were: Burukan in Siberia (because it had one of the worst records of daily RH > 100%), Saskatoon in Canada (because it was at a similar latitude to Burukan but at the opposite end of the longitude), Chicago, USA (because it was in a more temperate climate), New Delhi (as a subtropical climate) and Singapore (in the tropics).

Results

Results 1. 10/30 year monthly averaged data (Method 1)

Initial Scoping of the data comparing 30 year averages of Tmin, Tdew and RH (Tmin) for EWEMBI, W5E5 and UKesm(ISIMIP3bV2) and UKesm(ISIMIP3bV3).

Five sets of results were obtained. Note that the data in Table 2 is for individual grid cells with many grid cells having this high RH(Tmin) for more than 1 month. The EWEMBI1, W5E5 and UKesmV2 is the average for 1981-2010 while the GFDL4 and UKesmV3 is for 2001-2010 average. Note that all cells with Tmin<1 C have been removed, leaving a 68939 cells where RH(Tmin) > 110% for at least one month

Table 2: Number of cells where RH(Tmin) > 110% for at least one month where Tmin>1C from various ISIMIP models. n = 68939

Number of cells	EWEMBI1	W5E5	UKesmV2	GFDL4V3	UKesm1V3
RH(Tmin) > 110%	17159 (32%)	14049 (26%)	14313 (27%)	8513 (12%)	13808 (23%)
RH(Tmin) > 130%	420	329	349	208	367
RH(Tmin) > 160%	0	5	9	6	6

The RH(Tmin)>160% cells were located on a map and were all in the highlands of New Guinea where the altitude was sometimes up to 3000m. We are not at all confident with calculations at high altitude so we chose to exclude the high altitude cells. When these very high altitude cells were removed there were still a large percentage of cells with RH over 120% and 130% RH (Table 3).

Table 3: Number of cells where RH(Tmin)>110% for at least one month where Tmin>1C & altitude<1000m for various ISIMIP models. As the number of cells with Tmin>1C varied from model to model n varied between 50,000 and 60,000

Number of cells	EWEMBI	W5E5	UKesmV2	GFDL4V3	UKesmV3 from RH	UKesmV3 from SH
RH(Tmin) > 110%	13184 (22%)	10905 (19%)	11135 (20%)	11166 (22%)	10621 (21%)	11998 (24%)
RH(Tmin) > 120%	3595	1161	1237	229	243	532
RH(Tmin) > 130%	245	67	77	3	26	39

Comparison of weather station to grid cell data for Tmin, Tdew when RH(Tmin)>110%

In this section the restriction that Tmin>1C remains but the altitude restriction now becomes (for both W5E5 and the weather stations) that altitude < 400m. We use this more stringent restriction for altitude because for higher altitudes, weather station altitude and average grid cell altitude can be considerably different (see methods section).

Table 4 is a summary of the gridded data and weather stations showing the percentage of cells/stations where for at least one month, the 10-year monthly RH(Tmin) > 110%.

Table 4: Percentage of grid cells or weather stations where RH(Tmin) >110%

Model	RH(Tmin)>110%
W5E5	16%
EWEMBI	19%
UKesmV3	17%
Weather Stations	8%

One of the questions asked by PIK was whether the reason RH>100% was due to too low Tmin values or too high Tdew values. For this study we compared Tdew with Tmin for grid cells that were problematic (i.e. where RH(Tmin)>110%). We allowed for a possible +/-1C in Tmean and Tdew and only looked at those cells where the difference was greater than that. For this study we removed weather stations where RH(Tmin) > 110% to remove this confounding effect when comparing with W5E5 data. Table 5 gives the W5E5 data (which is similar to UKesm) compared to weather stations in the same grid cell. Number of cells = 98

Table 5: Comparison of both Tmin and Tdew for W5E5 data with good weather station data

	Minimum temperature	Dew Point Temperature
W5E5 = WS to +/- 1C	48%	51%
W5E5 < WS by 1C	41%	5%
W5E5 > WS by 1C	11%	44%

Note that RH(Tmin) > 110% occurs when Tdew > 1.3* Tmin.

Clearly the percentage of cells where Tmin in W5E5 is too low (giving a RH(Tmin)> 100%) is about the same as the percentage of cells where Tdew in W5E5 is too high (again giving a RH(Tmin) > 100%).

To further identify whether Tmin or Tdew were "faulty" we compared both these variables between weather stations and W5E5 by scatter graphs shown Figure 1 and 2.   The is no apparent difference between W5E5 and weather station data for Tmin or Tdew.

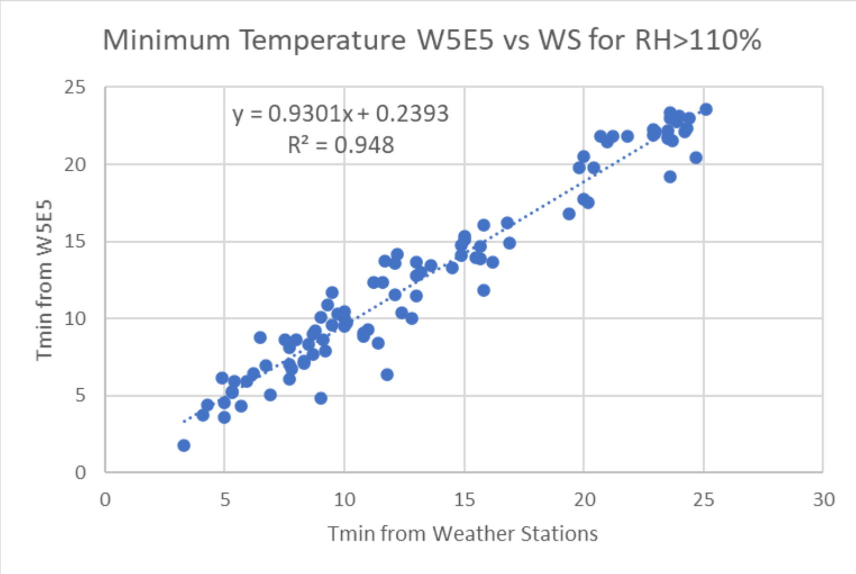

Figure 1 Comparison of weather station Tmin with W5E5 Tmin where RH(Tmin)> 110%

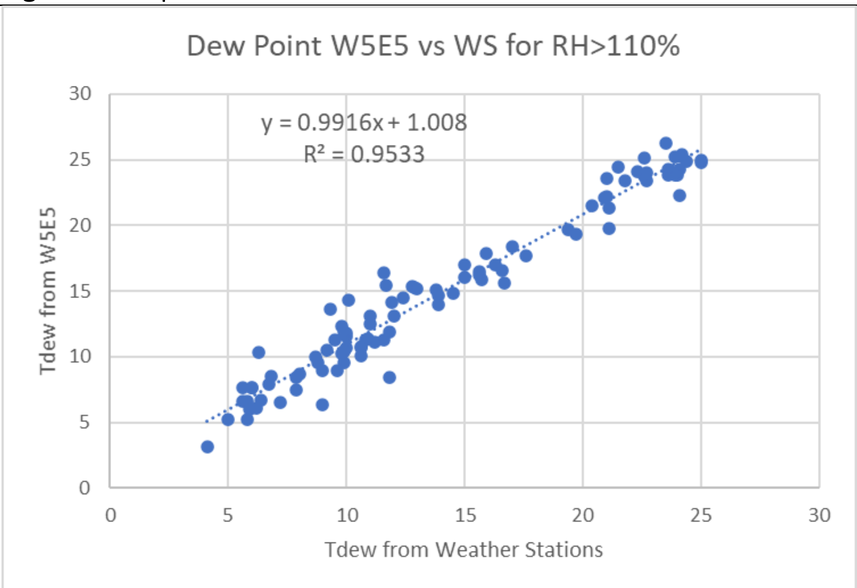

Figure 2 Comparison of weather station Tdew with W5E5 Tdew where RH(Tmin)> 110%

Results 2: Detailed daily and hourly weather station data (Method 2)

As these results showed nothing definitive, we decided to study some of the weather stations with the highest RH values in more detail by looking at the hourly (or three hourly) data as well as the daily data. For the hourly or 3 hourly data RH was **never** greater than 100% for any hour. To remove the possibility that this issue was cause by time zone differences, the Tdew of the day **before** and the day **after** the Tmin recording were tested to see if this made any difference to the daily Tdew>Tmin combination. In all 5 weather stations studied (see methods 2) the difference between the hourly data and the daily data had the lowest rms error on the same day when both Tmin and Tdew were used to calculate RH at Tmin. So the day before or day after measurement was not the problem. Closely looking at the hourly/three hourly data made it obvious why hourly data met the criteria of RH always less than 100% but the daily data did not meet that criteria (see table 6).

Table 6. A consecutive 24 hour UTC and local time recording of Tmin and Tdew (and RH calculated from Tmin and Tdew) in July 2020 from a reliable large airport weather station JG Diefenbaker International Airport, Saskatoon, Canada. This clearly illustrated the issue where hourly Tdew is never greater than Tmin, but in this case the daily average Tdew based on UTC time is 14.0C and based on Local time is 13.9C while Tmin is less than both of these at 11C. This particular day was chosen because it had one of the highest daily RH(Tmin) for temperatures>1C. This will be discussed in the next section.

Day	UTC Hour	Local Hour	Temp	Tdew	RH	Comment
21/20	0	18	22	15	65%	Start of UTC day
21/20	1	19	20	16	78%	Tdew relatively constant
21/20	2	20	20	16	78%	Falling temperature
21/20	3	21	18	15	83%	
21/20	4	22	17	15	88%	
21/20	5	23	14	14	100%	100% RH: Tmin reaches Tdew
21	6	0	14	13	94%	Start of local day
21	7	1	12	12	100%	100% humidity; Tdew dragged down by Tmin as water vapour condenses as dew
21	8	2	13	13	100%	
21	9	3	12	12	100%	
21	10	4	11	11	100%	Minimum Temperature
21	11	5	12	12	100%	
21	12	6	12	12	100%	
21	13	7	14	14	100%	
21	14	8	16	16	100%	Morning and sun rises increasing air temperature causing some of the dew to evaporate
21	15	9	19	15	78%	
21	16	10	22	15	66%	
21	17	11	23	16	65%	
21	18	12	22	14	61%	
21	19	13	24	14	54%	Tdew relatively constant
21	20	14	25	14	51%	
21	21	15	26	14	48%	Maximum temperature
21	22	16	25	14	51%	
21	23	17	24	15	57%	Average Tdew for UTC day = 14.04C
22/21	0	18	24	14	54%	Tdew relatively constant And similar to previous day
22/21	1	19	24	15	57%	
22/21	2	20	23	15	61%	
22/21	3	21	21	14	64%	
22/21	4	22	21	15	69%	
22/21	5	23	20	15	73%	Average Tdew for local day = 13.92C

Of note is the constant Tdew over time till the Tmin approaches Tdew which then drags the Tdew down with it. At no time is the RH(Tmin) greater than 100% - that is, Tdew is never greater than Tmin for this or any of the hourly data studied from the 5 weather stations in 2020 that all had daily RH(Tmin)>110%. The GSOD values for that day in Saskatoon were Tmin = 11.0C, Tdew =14.05C. Note the close match between

Tmin (11.0C) and Tdew (14.04C) for UTC time. Testing other days and other weather stations seems to suggest that daily Tmin and Tdew is based on a UTC day rather than a local day.

Discussion

It should be noted that there was a slight improvement in the number of grid cells where RH>110% (table 2 and 3) in going from ISIMIPV2 (June 2020 version) to ISIMIPV3 (May 2021 version), and that these both had substantially the same number of cells with RH(Tmin)>110% as W5E5 and less than the EWEMBI. It is apparent from the data in table 2 that the issue of RH(Tmin)>110% originates from W5E5 (and EWEMBI) and is NOT a problem of the bias correction method used.

The research then focussed on W5E5 to try to understand why the daily RH(Tmin) > 110% and whether this was caused by an "error" in Tmin or Tdew (as calculated from RH). Our studies showed approximately the same difference (table 5 and figures 1 and 2) between the W5E5 Tdew and Tmin and that of the weather station Tdew and Tmin. This led to a more detailed study of weather station data.

It was noted that daily weather station data also contained a large number of stations (8%) where RH(Tmin) >110% (see table 4). This took the issue away from the W5E5 (and EWEMBI) data and suggested the problem lay in daily weather station data that was used to construct W5E5 (and EWEMBI) data.

Detailed analysis of hourly data from 5 weather stations where RH(Tmin)>110%, an example of which is shown for 30 hours in table 6, all 5 weather stations showed the un-physical condition of RH>110% in daily weather station data but NOT in hourly (or 3 hourly) weather station data.

We believe that the cause of this problem in DAILY data is because of the averaging process when Tdew is dragged down by Tmin for part of the day. Using the data in table 6, Tdew stays relatively constant during daylight hours where it is well below Tmin. Then at midnight local time, Tmin approaches Tdew causing the Tdew to fall. Tdew decreases because water vapour is removed from the air by condensation triggered by the low Tmin. As it gets colder still, Tdew continues to fall. The data in table 6 has Tmin = 11C at 4am local time, but when Tdew is averaged for the day the Tdew is weighted to the daytime values which are higher than at Tmin. Hence the average value of Tdew is 14C, which when using daily values indicates that the RH>100% at Tmin.

Conclusion

The non-physical condition where RH > 100% is a problem in W5E5 and EWEMBI and models bias corrected from that data. The problem also exists, but to a lesser extent in weather station data. The most likely cause of the problem is that while the dew point temperature is relatively constant during the day, sometimes the minimum temperature drags the dew point down during the night while still maintaining the hourly RH=<100% criteria. However, the average of Tdew (taken over 24 hours) includes many hours where the Tdew is much higher before it is dragged down, while the Tmin time is a single point in the day. More rapidly cooling in a short period, lowers Tmin to be much lower (i.e. 5C lower) than the average of the daily Tdew. It seems the 24-hr mean formula approach for RH, SH or Tdew is not working well for places and times when Tmin reaches Tdew. The amount of moisture in the air near ground level is not constant over 24 hours as water condenses and evaporates over that time. While the moisture content is fairly constant from mid-morning to evening the difference between night-time and day-time moisture content leads to these apparently un-physical conditions of RH>100% based on daily averaged RH.

This would be possible to correct if we had 3 sets of RH, SH or Tdew data over the course of a day: for example: RHmin (at Tmax), RHmean (at Tmean) and RHmax (at Tmin). This data could only be obtained from weather station data that recorded accurate hourly data.

References:

Stefan Lange (2019) Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0) Geosci. Model Dev., 12, 3055–3070, 2019

Mark G Lawrence (2005) The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air A Simple Conversion and Applications. American Meteorological Society Feb 2005 p225

M Otto 2020 Developer of Hothaps-Soft available from climatechip.org/Analysis-tools. Updated 2020.

CRU (2021) Climate Research Unit: High-resolution gridded datasets (and derived products). Univ of East Anglia website crudata.uea.ac.uk/cru/data/hrg/ accessed 2021

Copernicus (2021) ERA5 climate re-analysis data at a single levels climate.copernicus.eu/climate-reanalysis accessed 2021.

Vaisala (2021) Make your job easier with humidity conversion formulas Vaisala Oyj web site: www.vaisala.com/en/lp/make-your-job-easier-humidity-conversion-formulas) accessed 2021.

NOAA (2021) National Weather Service, Weather Prediction Center, Meteorological Conversions and Calculations: www.wpc.ncep.noaa.gov/html/dewrh.shtml accessed 2021

Omni Calculator (2021) Dew Point Calculator www.omnicalculator.com/physics/dew-point accessed 2021.

NOAA GSOD (2021) Global Summary of the Day Weather Station Data data.noaa.gov/dataset/dataset/global-surface-summary-of-the-day-gsod accessed 2021.

NOAA ISH (2021) National Centres for Environmental Information Integrated Summary of the Day Global Hourly www.ncdc.noaa.gov/isd/data-access accessed 2021.