Evaluation of satellite-derived agro-climate variables in the Northern Great Plains of the United States

Rebecca Lemons a, Andrea Hewitt a, Gehendra Kharel a, Cherie New a, Andrei Kirilenko a & Xiaodong Zhang a
a UND, ESSP, Clifford Hall Room, 4149 University Avenue Stop 9011, Grand Forks, ND, 58202, USA
Accepted author version posted online: 05 Jan 2012. Version of record first published: 01 Feb 2012.

To cite this article: Rebecca Lemons, Andrea Hewitt, Gehendra Kharel, Cherie New, Andrei Kirilenko & Xiaodong Zhang (2012): Evaluation of satellite-derived agro-climate variables in the Northern Great Plains of the United States, Geocarto International, 27:8, 613-626
To link to this article: http://dx.doi.org/10.1080/10106049.2011.653408

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Evaluation of satellite-derived agro-climate variables in the Northern Great Plains of the United States

Rebecca Lemons*, Andrea Hewitt, Gehendra Kharel, Cherie New, Andrei Kirilenko and Xiaodong Zhang

UND, ESSP, Clifford Hall Room, 4149 University Avenue Stop 9011, Grand Forks, ND 58202, USA

(Received 4 August 2011; final version received 22 December 2011)

The climate of the United States Northern Great Plains region is highly variable. Modelling of agriculture in this region and similar locations depends on the availability and quality of satellite and ground data for agro-climate variables. We evaluated tropical rainfall measuring mission (TRMM) multi-satellite preparation analysis (TMPA) precipitation, atmospheric infrared sounder (AIRS) surface air temperature, and AIRS relative air humidity (RH). A significant bias was found within the temperature and RH products and no bias but an insufficient rain event detection skill in the precipitation product (probability of detection $\approx 0.3$). A linear correction of the temperature product removed the bias as well as lowered the root mean square deviation (RMSD). The bias-corrections for RH led to increased RMSD or worse correlation. For precipitation, the correlation between the satellite product and ground data improved if cumulative precipitation or only precipitation during the growing season was used.

Keywords: agro-climate; AIRS; precipitation; relative humidity; temperature; TMPA

1. Introduction

Although contemporary industrial agriculture is based on highly intensive management, it is still vulnerable to periodic crop failures due to weather variability. Multiple crop models have been developed to project past and future yields as a function of specific weather conditions, e.g. to trigger early food security alert systems (Verdin and Klaver 2002), to plan for future changes in agro-climate (Alcamo et al. 2007), or to estimate the effects of adaptation measures (Tubiello and Fischer 2007). Modelling crop growth requires the knowledge of various weather parameters, including temperature, precipitation and air humidity at the location of the field.

The amount of precipitation has an immediate impact on the type of crop that can be grown, the rates of photosynthesis, pollination, etc. (Klages 1942). Excessive or inadequate precipitation diminishes yields and degrades cropland (Chang 1968, Arriaga and Lowery 2003). Temperature controls the photosynthesis and also determines the length of the growing season, especially in the areas with sufficient year round precipitation (Chang 1968, Klages 1942). However, extreme temperatures suppress photosynthesis activity, impede growth rates and increase crop water

*Corresponding author. Email: Rebecca.Lemons@und.edu

ISSN 1010-6049 print/ISSN 1752-0762 online
© 2012 Taylor & Francis
http://dx.doi.org/10.1080/10106049.2011.653408
http://www.tandfonline.com
requirements (Biscoe and Gallagher 1977, Powell and Thorpe 1977). The amount of moisture in the air directly affects water balance. High potential evapotranspiration in low relative humidity (RH) and high air temperature conditions can trigger water stress in crops even when there is sufficient water within the soil (Chang 1968, Klages 1942). Crops often thrive under high RH conditions due to higher moisture availability (Breazeale et al. 1950, Breazeale and McGeorge 1953). However, exposure to the humid environments over extended periods of time may hamper transpiration rates, suppress nutrient uptake and increase susceptibility to disease. For these reasons, certain crops such as wheat, barley, alfalfa and clover are not suitable to grow in areas with high humidity and temperatures (Klages 1942). Altogether these agro-climatic variables dictate the types of crops that can be grown in a particular territory as well as put limits on the expected yield for the season (Biscoe and Gallagher 1977, Klages 1942).

To provide weather information that is adequate for agricultural modelling requires a dense network of meteorological stations that are well maintained. While this is typically not an issue for developed countries, developing countries tend to have fewer weather stations with shorter measurement records of lesser quality. In addition, agriculture in developing countries is primarily rain fed (95% of agriculture in sub-Saharan Africa and 90% in Latin America – Wani et al. 2009), which heightens the need for reliable weather data in these areas. The pressure to estimate future yield in the regions with sparse and less reliable ground measurements can be alleviated with satellite-based products, such as the United States Geological Survey water requirement satisfaction index yield model (Verdin and Klaver 2002), employed in the United States Department of Agriculture Famine Early Warning Systems Network food security outlook (http://www.fews.net). Remotely observed data from satellites provides an alternative that is better in spatial coverage and comparable in accuracy, if well validated.

Satellites have been used to estimate rainfall since the 1970s with infrared and microwave sensors; the latter are particularly useful because of their ability to penetrate clouds. In 1997, a joint US–Japan Tropical Rainfall Measuring Mission (TRMM) started with the goal of monitoring the tropical and subtropical precipitation (Ebert et al. 2007). In 2002, the atmospheric infrared sounder (AIRS) was launched to measure a suite of geophysical variables including temperature and humidity at various altitudes in the atmosphere. So far the efforts for validation have been primarily focused on vertical profiles, but few studies concentrated on evaluation of the variables at or near the surface, which is critical for agriculture production and decision making (Fetzer et al. 2003a, 2003b, Buehler et al. 2004, John and Buehler 2005, Moradi et al. 2010).

The Northern Great Plains of the United States is famously known for its weather extremes, both intra- and inter-seasonal (Karl et al. 2009), yet it remains one of the most productive agricultural regions in the world (Diamond 2002). The region has a thriving agriculture based economy that is driven by its ability to adapt to changes within the local climatic conditions (Saarinen 1966). Satellite observation offers a valuable alternative to the traditional weather stations network in monitoring and quantifying climatic change and its variability and therefore provides critical inputs to crop modelling and predictions. The goal of this study is to evaluate the performance of satellite derived precipitation, surface temperature and RH in this region and, if possible, to correct the bias. Such validation, conducted in an area with intensive agriculture and a developed network of weather stations, may provide
to insight on the quality of the satellite data and their usability in other locations, where the ground weather stations are sparse or not available.

2. Data and methodology

2.1 Site location

The US Northern Great Plains span across North Dakota, South Dakota, Minnesota, Wyoming, Montana and Nebraska, most of which are associated with grasslands habitats (Padbury et al. 2002). The elevation of this region ranges from \(~200\) m in the east to \(~4000\) m in the west (Figure 1). Depending on climate and soil, the region can be roughly divided into three agricultural zones. The eastern part of the region with more humid climate and deep mollisols, originally dominated by the tall-grass prairies has become one of the most productive agricultural areas in the world (Padbury et al. 2002). The major crops grown within this region include alfalfa, sugar beet, corn, oat, sunflower, canola, flax, potatoes, sorghum, soy beans, hay and spring wheat (Padbury et al. 2002, USDA-NASS 2008), and produces 27% (86,725,000 acres) of the total crops planted in the United States (USDA-NASS 2011). The western part is relatively dry and dominated by short-grass prairies, mainly used for ranching (58% of the land area) with only about 18% of the area used for crop production. Drought tolerant crops such as red spring wheat must be used in this region due to low and often variable rainfall (Padbury et al. 2002). Mixed prairies form a transitional zone between the western and eastern regions. Overall, the agriculture in this region is highly influenced by weather variability: extremely cold long dry winters, short warm summers, strong winds and constantly varying precipitation (Padbury et al. 2002).

2.2 Satellite data

NASA surface air temperature and RH products use the data collected from the AIRS onboard the Aqua satellite, which can be downloaded from the NASA website (http://disc.sci.gsfc.nasa.gov/AIRS/data-holdings/by-access-method). The Aqua satellite passes the region twice per day at 01:30 local time on the descending orbit.
orbit and at 13:30 ascending (Braverman et al. 2007). AIRS measures outgoing radiation in 2378 spectral channels in the infrared, near infrared and microwave wavelengths, from which three-dimensional maps of the air and surface temperature, water vapour, cloud formation and trace greenhouse gases like ozone, carbon monoxide, carbon dioxide and methane can be derived (Aumann et al. 2003). We evaluated the AIRS v. 5 level 3 products: surface air temperature and RH, distributed over a $0.5^\circ \times 0.5^\circ$ latitude and longitude grid (Parkinson and Greenstone 2000). Surface air temperature was estimated from the atmospheric temperature profile by extrapolating the local surface approximate reference height of 2 m (Jones et al. 2010). The RH data were from the lowest layer of the vertical profile, which is representative of the RH within the 92.5–100 kPa pressure levels (roughly 0–2 km) (Olsen 2007, Bedka et al. 2010). Satellite precipitation data were extracted from NASA Goddard Space Flight Center TRMM multi-satellite preparation analysis (TMPA). For this study, we used the 3B42 v. 6 product, which combines the measurements from multiple satellites with the rain gauge data and is available over a global grid of $0.25^\circ \times 0.25^\circ$ between 50°N and 50°S with 3 h temporal resolution (Huffman et al. 2007).

### 2.3 Ground data

Three different collections of ground data were used in this study (Figure 1). United States historical climatology network (USHCN) maintains the data from 1218 meteorological stations across the lower 48 states with rigorous quality control on the data, which include maximum and minimum temperature, precipitation, snowfall and snow depth, all reported daily and monthly. For precipitation and temperature, we used daily data from 98 meteorological stations within the Northern Great Plains region from 1 January 2003 to 31 December 2009. The maximum temperature was compared with the AIRS daytime data (13:30 local time) and the minimum temperature with night-time data (1:30 local time). The daily mean temperature based on meteorological data was derived from the mean of maximum and minimum temperatures and that for satellite from the mean of AIRS daytime and night-time readings.

Since the USHCN data do not include RH, we used two additional meteorological networks, North Dakota agricultural weather network (NDAWN) and AgriMet. NDAWN has been in operation since 1989 and includes 81 stations scattered throughout North Dakota and the surrounding states. Each station measures hourly air and soil temperature, RH, wind speed and direction, solar radiation, rainfall, barometric pressure, dew point and wind chill. The AgriMet programme, which began in 1983, is part of the Pacific Northwest Cooperative Agricultural Weather Network and consists of 70 agricultural weather ground stations. These stations measure air temperature, precipitation, solar radiation, soil temperature, dew point, RH, barometric pressure and wind speed and direction every 15 min. Only 10 AgriMet stations are located within our study area. Among a total of 91 NDAWN and AgriMet stations located within the region, 72 had corresponding AIRS RH data that passed NASA quality control. The additional 19 stations within the study area contained corresponding AIRS data, which did not pass NASA quality control and was therefore not released as level 3 data or were outside of the coverage area for AIRS that day. For comparison with AIRS daytime and night-time records, four 15-min RH values from ground stations, one set from 01:00–02:00 local time and the other from 13:00–14:00 local time were averaged.
2.4 Satellite and ground data comparison
For each cell (0.5° × 0.5° for the surface air temperature and RH; 0.25° × 0.25° for precipitation) of the gridded satellite product, the ground stations located within the cell were selected for comparison. If no ground data were found within the cell, the corresponding satellite data were discarded from further consideration; if more than one station was located within a grid, then the data were averaged among the stations. The comparison between satellite and ground measurements was evaluated using the Pearson correlation (\( \rho \)), the root mean square deviation (RMSD) and the mean bias (MB). The statistical analysis was also done per station and growing verses non-growing season for all the products and monthly mean for precipitation only. Additional analyses were performed for precipitation (Ebert et al. 2007, Vila et al. 2009, Tobin and Bennett 2010), including estimates of four performance indicators: the frequency bias index (FBI), false alarm ratio (FAR), probability of detection (POD) and threat score (TS; Figure 2).

3. Results and discussion
The results of the comparison indicated that the performance of the remotely sensed data varied between the products (Table 1). Generally, both surface air temperatures of daytime and night-time orbits correlated well with the measured maximum and minimum daily temperatures (\( \rho > 0.90 \)) with an MB of −0.5°C and 5°C, respectively. The considerable difference between the daily minimum temperature and the satellite estimate (MB = 5°C) was likely due to the difference between the time of data acquisition (1:30 local time) and the coldest time of the day, typically at dawn. RMSD for daytime temperature (3.7°C) was lower than that for the night-time measurement (5.4°C), but both were higher than the reported AIRS retrieval goal of 2°C for surface air temperature (Gao et al. 2008). Tobin et al. (2006) also found that the AIRS temperature product over the Southern Great Plains (SGP) region did not meet AIRS performance requirements, in part due to the seasonal differences in land surface emissivity in the agricultural areas. The daily mean temperature estimated from AIRS showed an MB of 2.2°C as compared to the mean values estimated from the ground stations (Table 1). To correct this bias, we subtracted the value of 2.2°C from each daily mean temperature estimated from AIRS (Table 1). Correcting the

![Figure 2](image-url)

Figure 2. A 2 × 2 contingency table, modified from Wilks (1995). The cells represent the number of rain events detected by the satellite (a and b) and at the ground station (a and c). The frequency bias index (FBI) measures the over or under estimation of the number of rain events, false alarm ratio (FAR) measures the fraction of false rain detections, probability of detection (POD) measures the fraction of correctly detected rain events, and the threat score (TS) measures the overall skill of the satellite. The perfect scores for FBI, FAR, POD, and TS are 1, 0, 1, and 1 respectively.
bias decreased the overall RMSD of mean daily temperature from 4.2°C to 3.6°C (Table 1). Even though this is still higher than the 2°C RMSD error of the AIRS Team retrieval goals, applying the bias-correction allowed to reproduce the temperature distribution as observed by ground stations: using the Kolmogorov–Smirnov (K–S) test, there is no significant difference between distributions at 0.73 level (Figure 3a). However, the RMSD values for the corrected surface air temperature, ranging from 3.6 to 5.4°C (Table 1), are still higher than the AIRS temperature retrieval goal of 2°C, even though they are consistent with those found in a similar study conducted in China (Gao et al. 2008).

Table 1. Comparison of the satellite and ground daily precipitation measurements for the Northern Great Plains: root mean square deviation (RMSD), the Pearson’s correlation coefficient ($\rho$), and the mean bias (MB).

<table>
<thead>
<tr>
<th>Agro-climate variables</th>
<th>N</th>
<th>RMSD</th>
<th>$\rho$</th>
<th>MB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily precipitation</td>
<td>206,823</td>
<td>4.9</td>
<td>0.46</td>
<td>−0.03</td>
</tr>
<tr>
<td>Monthly precipitation</td>
<td>7,026</td>
<td>0.72</td>
<td>0.82</td>
<td>−0.03</td>
</tr>
<tr>
<td>Daytime temperature</td>
<td>149,335</td>
<td>3.7</td>
<td>0.92</td>
<td>−0.5</td>
</tr>
<tr>
<td>Night-time temperature</td>
<td>149,335</td>
<td>5.4</td>
<td>0.95</td>
<td>5.0</td>
</tr>
<tr>
<td>Mean temperature</td>
<td>149,335</td>
<td>4.2</td>
<td>0.96</td>
<td>2.2</td>
</tr>
<tr>
<td>Corrected mean temperature</td>
<td>149,335</td>
<td>3.6</td>
<td>0.96</td>
<td>0</td>
</tr>
<tr>
<td>Daytime RH</td>
<td>84,263</td>
<td>15.8</td>
<td>0.74</td>
<td>7.3</td>
</tr>
<tr>
<td>Night-time RH</td>
<td>85,675</td>
<td>21.8</td>
<td>0.25</td>
<td>13.1</td>
</tr>
</tbody>
</table>

Note: The units are mm/day (°C), and % for precipitation, temperature and relative humidity, respectively.

Figure 3. Frequency distribution of the ground and satellite data for temperature (a), precipitation (b), and RH daytime (c) and night-time (d).
The daytime RH satellite measurements compared better with the ground measurements than the night-time measurements, as evidenced by correlation and RMSD (Table 1), yet both day and night-time satellite RH data distribution was significantly different from the ground observations according to the K–S test (0.001 level – Figures 3(c) and (d)). Both daytime and night-time RH had a wet bias for lower RH values and a dry bias for higher RH values even though the hinge point was slightly different between the two (~60% for daytime and ~80% for night-time). A similar non-linear behaviour was found by Bedka et al. (2010), who reported wet biases in water vapour pressure at the atmospheric radiation measurement’s (ARM) SGP site during very dry events and a night-time dry bias during very wet events; note that these biases were not seen in the other two sites (The ARM Tropical Western Pacific (TWP) site and The ARM North Slope of Alaska (NSA) site) covered within that study. This could partly be due to the altitude mismatch between the satellite and the ground measurements while the AIRS RH product represents an average over a roughly 2-km air column above the ground, the weather stations typically measures the RH at a height of 2 m. Over the Great Plains, local weather and atmospheric stability can experience dramatic change at short time intervals, and therefore it is likely that a sample at 2 m differs from the mean value over a 2-km air column and this difference could change significantly over the day (Tobin et al. 2006).

Various studies (Buehler et al. 2004, Gettelman et al. 2010, Bedka et al. 2010, Moradi et al. 2010) have suggested that the satellite RH measurements are less reliable under elevated cloudiness. Fetzer et al. (2006) speculated that cold weather cloud formations could be the cause of an increase in bias in satellite water vapour products. To examine this cloud effect, we filtered out the satellite-measured RH for the days and nights with high cloudiness and found only a slight decrease in RMSD for daytime and a slight increase in RMSD for night-time comparisons (Figure 4). We also noticed that both satellite datasets contain erroneous values of RH up to 150% (1.5% of daytime data and 5% of night-time data), yet, removing these erroneous values did not improve the comparison (results not shown). We explored multiple methods to correct the bias in AIRS RH, however, all the corrections would

![Figure 4](image_url)

**Figure 4.** The effect of cloud cover on the estimate of RH from the AIRS observation in daytime (grey) and night-time (black). The RMSDs were calculated by progressively removing the satellite data measured with fractional cloud cover higher than the value on the x-axis.
increase RMSD and decrease correlation between the corrected satellite and ground data. This is likely due to a combination of heterogeneous and rapidly changing humidity distribution found in the Northern Great Plains region and the difference in reference heights between the AIRS satellite RH product (mean value for 0–2 km) and the ground observation (~2 m).

We did not confirm previous studies that connect possible comparison issues of the satellite RH estimation to high cloudiness interference. If the AIRS RH data were to be extrapolated to the surface level, similar to AIRS surface temperature data, the differences seen between AIRS and ground stations in this study might be further constrained. Due to these unknown variables, no bias-correction was applied to the AIRS RH data. Also in the following analysis, we focused only on daytime RH measurements because of the low correlation for the night-time data.

Satellite estimates of daily precipitation showed virtually no bias as compared with the ground measurements, which was expected given that TMPA estimates were already bias-corrected (Huffman et al. 2007). In terms of distribution, overall good agreement was found between the TMPA and ground observations at various precipitation intensities (Figure 3d), which was also supported by FBI scores (Table 2), yet both K–S and Mann–Witney tests showed that the distributions were significantly different at 0.1 level, but not at 0.05 level. The skill of TMPA was found very poor in terms of FAR, POD and TS scores (Table 2): only one-third of the rainfall events were detected successfully (POD = 0.21–0.34) while about 70% were detected as false-alarms (FAR = 0.68–0.78). Also, the overall fraction of correctly diagnosed events by TMPA was very low (TS = 0.12–0.19). It was found that the performance of TMPA deteriorated from lighter to heavier rainfall events as indicated by the increase in FAR and the decreases in both POD and TS.

The monthly averaged daily precipitation showed improved RMSD (0.71 mm/day vs. 4.9 mm/day for daily data) and correlation with ground observations (0.82 vs. 0.46). This is not surprising, because the inherent uncertainty in the daily product was lowered through averaging in the monthly product (Scheel et al. 2010). The poor correlation of daily precipitation could possibly be due to the inherent random errors in rain gauge measurements (Ciach 2003) and due to the nature of TMPA algorithm, which includes an adjustment of the satellite-derived estimations of precipitation to match the rain-gauge measurements at a monthly scale (Huffman et al. 2007), which may not improve the daily or hourly estimates (Scheel et al. 2010). The poor performance of TMPA daily

<table>
<thead>
<tr>
<th>Precipitation (mm/day)</th>
<th>FBI</th>
<th>FAR</th>
<th>POD</th>
<th>TS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entire dataset</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5–10</td>
<td>1.07</td>
<td>0.68</td>
<td>0.34</td>
<td>0.19</td>
</tr>
<tr>
<td>10–30</td>
<td>0.90</td>
<td>0.72</td>
<td>0.25</td>
<td>0.15</td>
</tr>
<tr>
<td>&gt;30</td>
<td>0.96</td>
<td>0.78</td>
<td>0.21</td>
<td>0.12</td>
</tr>
<tr>
<td>Growing season</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5–10</td>
<td>1.10</td>
<td>0.64</td>
<td>0.40</td>
<td>0.23</td>
</tr>
<tr>
<td>10–30</td>
<td>0.88</td>
<td>0.69</td>
<td>0.27</td>
<td>0.17</td>
</tr>
<tr>
<td>&gt;30</td>
<td>0.95</td>
<td>0.77</td>
<td>0.22</td>
<td>0.13</td>
</tr>
<tr>
<td>Non-growing season</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5–10</td>
<td>1.01</td>
<td>0.78</td>
<td>0.22</td>
<td>0.13</td>
</tr>
<tr>
<td>10–30</td>
<td>1.01</td>
<td>0.86</td>
<td>0.15</td>
<td>0.08</td>
</tr>
<tr>
<td>&gt;30</td>
<td>0.98</td>
<td>0.93</td>
<td>0.07</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Note: See Figure 2 for explanation of the frequency bias index (FBI), false alarm ratio (FAR), probability of detection (POD) and threat score (TS).
precipitation could be due to errors at the sensors and product algorithm (Scheel et al. 2010). For example, the global algorithm of infrared data processing does not consider the distance between the cloud top and earth’s surface. Since the precipitation is derived from the brightness temperature at cloud top (Levizzani et al. 2002), excluding altitude information in the infrared data processing could affect TMPA estimates. Further, microwave signals are dependent on surface types (Ferraro et al. 1998). Cold land surfaces, and ice or snow covered areas could make estimation difficult due to strong scattering (Huffman and Bolvin 2011).

The performance of satellite products also showed a certain spatial variability, as depicted in Figure 5 showing RMSD for each ground station. Comparison of

![Figure 5](image_url)

Figure 5. Comparison between the satellite and ground measurements: spatial variability of RMSD for corrected AIRS temperature (a), relative humidity (b), and precipitation (c). Isotherms (a), isohumes (b) and isohyets (c) are based on the ground data.
Figures 1 and 5(a) suggest that the bias-corrected AIRS mean temperature data has the highest RMSD errors at the higher elevations. This was consistent with Gao et al. (2008) who evaluated the AIRS surface air temperature in China and found that high RMSD errors were located at high terrain. The RMSD for daytime RH decreased from east to west (Figure 5(b)), which might be an artefact due to a general scarcity of satellite coverage in the western part of the Northern Great Plains. Despite this, the pattern approximately matched the overall east to west RH yearly mean gradient in the Northern Great Plains. For daily precipitation, RMSD values also decreased from east to west across the region, possibly due to the effect of the precipitation gradient (Figure 5(c)).

Besides spatial differences, there are also significant seasonal variations in the satellite data quality. Corrected mean satellite and ground temperature showed virtually the same correlation for both growing and non-growing season with notable difference in RMSD between the seasons (3.3°C vs. 3.9°C) (Table 3). The difference in RMSD could be attributed to landscape properties that change with season. In the agricultural region of the Northern Great Plains, non-growing season is likely to have bare soils and little to no vegetative cover, while growing season will have new growth of various vegetation types. The substantial differences in surface emissivity between bare soil and vegetation land covers would affect the retrieval of air temperature from AIRS measurements, producing larger errors over the Northern Great Plains region (Prihodko and Goward 1997, Parkinson and Greenstone 2000).

Higher correlations between the satellite and ground data were observed during the growing season than the non-growing season for both precipitation and RH. However, RMSD for non-growing season data appeared to be lower than that for the growing season for precipitation (Table 3), possibly due to seasonal distribution of precipitation, with lower amounts of precipitation during the cold season. At the same time, relatively better FAR, POD and TS scores for the growing season (Table 2), consistent with earlier studies (Scheel et al. 2010), suggest better performance of TMPA in the growing season.

It would be beyond the scope of this article to report the evaluation of yield predictions for different crops using the remotely sensed data that we have examined above. However, it is of interest to show how sensitive a crop model might be to these errors. We ran a sensitivity analysis, using the CropGro wheat model, a

<table>
<thead>
<tr>
<th>Climate variable</th>
<th>Growing season</th>
<th>Non-growing season</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily precipitation</td>
<td>118,408</td>
<td>88,415</td>
</tr>
<tr>
<td>Monthly precipitation</td>
<td>3,994</td>
<td>3,032</td>
</tr>
<tr>
<td>Mean temperature</td>
<td>88,024</td>
<td>61,311</td>
</tr>
<tr>
<td>Corrected mean temperature</td>
<td>88,024</td>
<td>61,311</td>
</tr>
<tr>
<td>Daytime RH</td>
<td>50,399</td>
<td>33,864</td>
</tr>
</tbody>
</table>
member in the DSSAT-4 (http://www.icasa.net/) suite of yield models. The model parameters were calibrated to simulate the spring wheat production in North Dakota for the current climate. We selected eight locations throughout the state to take into account the existing strong West-East precipitation gradient, the North-South temperature gradient and soil variability. For each location, we ran the model once using the current climate to predict the ‘true’ yield. Then we perturbed one of the input agro-climatic variables (air temperature, precipitation or RH) with random noises generated following the distribution characterized in Table 1. For example, to simulate how the use of AIRS air temperature without the bias correction would affect the yield prediction, the current air temperature was perturbed by a normally distributed random noise with a mean value of 2.2°C (MB) and standard deviation of 3.6°C (RMSE). This ‘perturbed’ process was repeated 10 times for each variable and the results were compared with the ‘true’ yield. The mean values of these 10 comparisons were reported in Table 4.

The spring wheat model is very sensitive to the bias in air temperature estimates, causing an average of 10% error in prediction, probably due to higher sensitivity of model to soil moisture variations under dry climate. With bias-corrected temperature, the sensitivity is largely constrained with an average error of 3%. The sensitivity to precipitation errors is about 2%, probably because of either or a combination of the following two reasons: (1) no bias was found in precipitation product; and (2) spring wheat is resistant to water stress. Finally, the errors in RH, though relatively large themselves as compared to other two variables, exert a negligible impact for this particular model (<1%). Here we have to stress that the sensitivity of crop model is model and crop specific, and we intend to report a comprehensive analysis of model sensitivity analysis in a different study.

It should be noted that this study has some limitations. First, we assumed that the traditional ground measurements represented the ‘true’ values of the meteorological variables, while numerous studies proved otherwise in certain conditions; for example, the rain gauge measurements performed with tipping buckets demonstrate significant biases during high wind and high precipitation events.

### Table 4. The sensitivity of spring wheat production model, CropGro, to errors in the satellite products of T, P and RH evaluated in eight locations in North Dakota.

<table>
<thead>
<tr>
<th>Name</th>
<th>Location</th>
<th>T + n (0,3.6)</th>
<th>T + n (2.2,4.2)</th>
<th>P + n (0,4.9)</th>
<th>RH + n (0,18.8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Langdon</td>
<td>NE</td>
<td>-18.9</td>
<td>-12.4</td>
<td>-5.5</td>
<td>-2.7</td>
</tr>
<tr>
<td>Fargo</td>
<td>SE</td>
<td>0.4</td>
<td>-8.8</td>
<td>-0.3</td>
<td>-1.0</td>
</tr>
<tr>
<td>Carrington</td>
<td>CE</td>
<td>-5.4</td>
<td>-12.9</td>
<td>0.3</td>
<td>-0.1</td>
</tr>
<tr>
<td>Minot</td>
<td>CC</td>
<td>-1.5</td>
<td>-18.4</td>
<td>-6.3</td>
<td>0.1</td>
</tr>
<tr>
<td>Streeter</td>
<td>SC</td>
<td>-1.9</td>
<td>-1.2</td>
<td>0.9</td>
<td>1.1</td>
</tr>
<tr>
<td>Williston</td>
<td>NW</td>
<td>5.3</td>
<td>-6.3</td>
<td>1.4</td>
<td>2.1</td>
</tr>
<tr>
<td>Dickinson</td>
<td>NC</td>
<td>-5.7</td>
<td>-10.5</td>
<td>-1.7</td>
<td>2.1</td>
</tr>
<tr>
<td>Hettinger</td>
<td>NS</td>
<td>1.4</td>
<td>-9.7</td>
<td>-5.5</td>
<td>1.9</td>
</tr>
<tr>
<td>Mean</td>
<td>–</td>
<td>-3.29</td>
<td>-10.03</td>
<td>-2.09</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Note: The sensitivity is evaluated as (‘perturbed’ – ‘true’)/‘true’, where ‘true’ denotes the predicted yield when the actual climatic variables of T, P and RH were used and ‘perturbed’ denotes the estimated yield when one of the climatic variables was perturbed with random noises \(n(m, \sigma)\) generated with a mean value of \(m\) and a standard deviation of \(\sigma\). The values of \(m\) and \(\sigma\) were assigned to those of MB and RMSE in Table 1 respectively to simulate the errors introduced by the use of satellite products. Also the perturbation was repeated 10 times for each location and the mean sensitivity of these ten repetitions is reported.
The USHCN data have already been corrected, however, the use of other sources of meteorological data such as the National Weather Service Next Generation Weather Radar (NEXRAD) may further improve the analysis. In our study, multiyear measurements over a large territory were used, which should have reduced the effect of any particular single data acquisition.

Second, an additional error that is inherent in almost all comparisons between satellite and ground measurements could be due to the mismatch of spatial scales between the two measurements, both horizontal and vertical. For example, in our study the satellite RH data represented the averaged humidity in a 2-km air column above ground, as opposed to the 2-m high ground data measured by the meteorological stations. Similarly, the cell sizes of satellite Level 3 products that we used were ~20 km for precipitation and ~45 km for temperature and RH, while typical range of the ground based measurements was of the order of 100 m.

4. Conclusion

The results of our evaluation of satellite temperature, precipitation, and RH for the Northern Great Plains region differ highly among the products. The AIRS surface temperature was found to have a varying bias as compared with the minimum, maximum and mean daily temperature. After correction, the product can generally reproduce the mean daily temperature distribution in the region. The bias-corrected temperature product can be used in applications with cumulative temperature inputs, such as growing degree days, but might not be suitable for applications requiring accurate daily temperature inputs because of its relatively large RMSD.

For precipitation, we found that the daily TMPA satellite and ground data were poorly correlated. However, there was good correspondence between the TMPA monthly precipitation and ground data, consistent with earlier studies (e.g. Su et al. 2008). Seasonal analysis showed better performance of TMPA in detecting precipitation events during growing season as compared to non-growing season. Based on our findings, we suggest that while the satellite product can be used to estimate spatial distribution of precipitation on a monthly basis, or to provide an input to a stochastic daily weather generator, the daily precipitation product cannot be reliably used.

Both AIRS RH products were found to have a bias which was hard to correct for because of the AIRS RH frequency distribution differs from the ground based observations in a non-linear manner. Overall, daytime RH compared better with ground observations than night-time RH. If AIRS satellite data were to be used in a model, we would recommend only using the daytime RH data.

In crop modelling, the bias-corrected data should be used. And for spring wheat, the residual errors found in the remote sensing products (i.e. RMSE) do not seem to affect the model performance very much, with the largest sensitivity due to air temperature (~3%). However, more tests are needed to fully evaluate the use of satellite products for crop modelling.

Acknowledgements

This article summarized the results of a graduate student research project, partially funded through NASA grant NNX10AH20G. We thank Dr. Edward Olsen from NASA JPL for his help with the AIRS RH product. We also would like to thank Dr. Hojin Kim, Ashley Tooley and Yingxi Shi from the University of North Dakota for their help in data processing.
References


