multi-year satellite estimates. This single satellite image will be used to nowcast DG production and to produce forecasts. We also present a novel method to parameterize the correlation between satellite pixels using the relative cloudiness between them.

We describe the satellite derived irradiance models and observation data in Section 2, the OI method in Section 3, and three ways to estimate covariances between locations in Section 4. These covariances are critical to the success or failure of OI. A method to correct for satellite geolocation errors is described in Section 5, and parameter tuning is detailed in Section 6. We present and discuss the results of applying OI to Tucson, AZ in Section 7. Finally, a summary of the work is provided in Section 8.

2. Models and observations

2.1. Satellite derived irradiance models

To investigate how well OI works with different types of prior information, we use two different models to convert satellite images to GHI maps. The resulting GHI maps are called the "background" or "prior" in OI and will be denoted by $\mathbf{x}_b$. Both models use images from the GOES-W geostationary satellite that cover the city of Tucson, AZ (roughly 75 × 80 km). An example of a visible albedo image derived from the visible channel of GOES-W is shown in Fig. 1.

One of the models is a physical model called the University of Arizona Solar Irradiance Based on Satellite (UASIBS) model (Kim et al., 2016). UASIBS uses the visible and infrared images from the GOES-W satellite to generate a cloud mask. Then, parameterized cloud properties determined from the infrared images are used in a radiative transfer model to determine the surface GHI. This GHI estimate has the same resolution as the visible channel of the GOES-W satellite (approximately 1 km).

The second model is a semi-empirical model, which we refer to as the SE model. This model is based on the SUNY model which applies a regression to the visible channel of the GOES-W satellite (Perez et al., 2002). The only differences between our SE model and the SUNY model are that the dynamic range is set with the 3 months of data used in this study instead of the recommended 60 day window with seasonal corrections and that the specular correction factor was neglected.

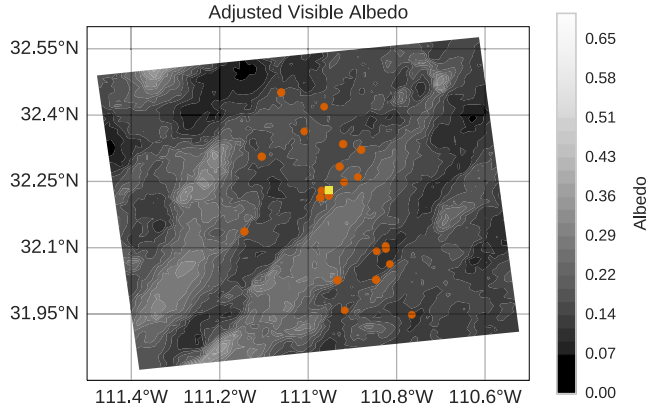

Fig. 1. Visible albedo image derived from the visible channel of the GOES-W satellite. Lighter colors indicate cloudier areas. The orange dots represent the locations of the sensors used in this study which includes both irradiance sensors and rooftop PV systems as described in Section 2.2. The yellow square in the center indicates the location of a calibrated GHI sensor on the University of Arizona campus. The image covers an area of roughly 75 × 80 km over Tucson, AZ. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

To remove effects of the diurnal cycle and ease incorporation of data from rooftop PV systems, all images were converted into clear-sky index images by dividing the estimated GHI by a clear-sky GHI estimate. The resulting values of clear-sky index range from nearly 0 for an overcast sky to 1 for a cloud-free sky. These conversion algorithms do not take into account image time-stamp inaccuracies or satellite geolocation errors, but corrections for those errors will be discussed in Section 5.

2.2. Ground observations

The observation data are collected from 22 sensors including a calibrated NREL MIDC sensor (Wilcox and Andreas, 2010), custom irradiance sensors (Lorenzo et al., 2014), and data from rooftop PV systems. The sensor locations are indicated by orange circles in Fig. 1.

Irradiance observations were averaged to 5 min to match PV data that are reported as 5 min averages. This averaging is consistent with the inherent averaging due to the satellite spatial resolution. We note that all data sources (ground sensors and satellite images) are available in near real-time so that the OI corrected GHI images can be used as a basis for forecasts or DG nowcasts.

All data were converted to clear-sky index data using clear-sky expectations for each sensor. To produce the clear-sky expectation for one day, the measurements from preceding clear days within one week are averaged to produce an initial estimate. This initial estimate is then scaled to match the clear times on the day of interest to account for differences in turbidity or temperature. This method simplifies the calculation of clear-sky expectations for the rooftop PV systems because no parameters about the system (directional response, peak power) are assumed. The clear-sky expectations and clear-sky index data was inspected manually to confirm the quality. Note that the ground observation data may experience cloud enhancement events which lead to clear-sky indices greater than 1.

We restrict our data and analysis to solar zenith angles less than 60°. At times, we also withhold sensors from the OI routine and use these sensors to validate how well OI performs for other locations in the image besides the input sensor locations.

2.3. Data set description

About 1300 satellite images collected over April, May, and June 2014 were converted to irradiance images with the two models and paired with the corresponding ground observations. We randomly divide the data set into a training set with 437 images (252 clear and 185 cloudy images) and a verification set of 874 images (504 clear and 370 cloudy images). The training set is used to tune parameters for OI as described in Section 6. The verification set is used for error analysis and to draw conclusions about OI.

The distinction between clear and cloudy satellite images will become important in Section 3 for determining sensor error variances. Clear times are identified using a combination of the UASIBS estimates and the ground sensor data. Specifically, if the minimum value of a UASIBS clear-sky index image is greater than 0.8, the mean of the image is greater than 0.99, and the second largest deviation from 1 of any of the ground observations is less than 0.05, then we classify the image as clear. This procedure accurately identifies times at which no clouds exist in the area of study. Other methods can also be used to perform this classification (Reno and Hansen, 2016; Escrig et al., 2013; Ghonima et al., 2012), but our simple method is sufficient for our purposes.