observation locations to locations throughout the background image. Similar to $\mathbf{R}$, $\mathbf{P}$ is defined such that

$$\mathbf{x}_b = \mathbf{x}_t + \mathbf{g}, \quad \mathbf{g} \sim N(\mathbf{0}, \mathbf{P}) \tag{4}$$

where $\mathbf{x}_t$ is the "true" value of the satellite derived clear-sky index image and $\mathbf{g}$ is a random vector sampled from a multivariate normal distribution with mean 0 and covariance $\mathbf{P}$.

We will now describe three methods to calculate $\mathbf{P}$:

1. *Empirical:* $\mathbf{P}$ calculated empirically from all of the background images,
2. *Spatial:* $\mathbf{P}$ with correlations parameterized based on the physical distance between pixels, and
3. *Cloudiness:* $\mathbf{P}$ with correlations parameterized based on the difference in cloudiness between each pixel.

4.1. Empirical covariance

The empirical $\mathbf{P}$ is calculated by assuming that satellite-derived clear-sky index images are sampled from the same multivariate normal distribution and then simply computing the covariance using all images in the training data set. This assumption is likely invalid given the high probability of clear days leading to a non-Gaussian distribution. An analysis computed with this type of $\mathbf{P}$ gives non-physical results, as described in Section 7, but is included for comparison.

4.2. Correlation matrix parameterization

Before describing spatial and cloudiness covariances, it is useful to decompose $\mathbf{P}$ into a diagonal variance matrix, $\mathbf{D}$, and a correlation matrix, $\mathbf{C}$ as

$$\mathbf{P} = \mathbf{D}^{1/2}\mathbf{C}\mathbf{D}^{1/2}. \tag{5}$$

Here, $\mathbf{D}$ sets the scale of the errors while $\mathbf{C}$ describes how errors and information spread. We obtain $\mathbf{D}$ in a similar manner as we do for $\mathbf{R}$, we use a number of clear images from the training data set to estimate the variance of each pixel in the background individually. The errors in $\mathbf{x}_b$ come mainly from the satellite image to ground irradiance conversion that often exhibits large differences in error between clear and cloudy images. Thus, we allow for a tunable scaling factor, d, in the construction of $\mathbf{D}$ for cloudy images to account for possible model error differences between clear and cloudy skies so that

$$\mathbf{D} = d\mathbf{D}' \tag{6}$$

where $\mathbf{D}'$ is the variance estimated from the clear images.

The correlation matrix $\mathbf{C}$ defines how information is transferred from the sensor locations to other locations in the satellite estimate. $\mathbf{C}$ can be parameterized based on the spatial distance between points in the background as in Ruiz-Arias et al. (2015) or, as we demonstrate, one might rely on information in the current satellite image, such as cloudiness.

To construct the elements of $\mathbf{C}$, c_{ij}, we apply a correlation function, k, to the distance metric r computed between each pixel i and j

$$c_{ij} = k(r_{ij}). \tag{7}$$

Any number of covariance functions, k, can be chosen; see Rasmussen and Williams (2005) for a partial list. We chose to study a piece-wise linear correlation function

$$k(r) = \begin{cases} 1 - \frac{r}{l} & r < l \\ 0 & r \geqslant l \end{cases}, \tag{8}$$

an exponential correlation function

$$k(r) = \exp\left(-\frac{r}{l}\right), \tag{9}$$

and a square exponential correlation function

$$k(r) = \exp\left(-\frac{r^2}{l^2}\right). \tag{10}$$

For each correlation function, l is a characteristic length that we tune with a training data set for each choice of k to minimize error as described later in Section 6.

4.3. Spatial covariance

The distance metric for the spatial correlation parameterization is the standard Euclidean distance (once locations are mapped to a two dimensional plane using a map projection),

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \tag{11}$$

Thus, the spatial covariance $\mathbf{P}$ is constructed by applying Eqs. (5)–(7) and (11) with a tuned k, l, and d as described in Section 6.

4.4. Cloudiness covariance

For what we call cloudiness covariance, we parameterize $\mathbf{C}$ based on the difference in cloudiness in the visible satellite image. This corresponds to only adjusting the cloudy areas with observations that are experiencing similarly cloudy sky and leaving the clear areas to be adjusted by observations of the clear sky. This adjustment is made without consideration of the spatial distance between pixels. We use the adjusted visible albedo calculated from the visible satellite image rather than the processed clear-sky index maps to compute the correlation. This avoids cloud representation errors that may arise in the satellite to irradiance conversion; for example, note how UASIBS fails to produce clouds in many areas of Fig. 2. Also note that because this parameterization depends on the visible satellite image, $\mathbf{C}$ and subsequently $\mathbf{P}$ are calculated for each image individually.

To calculate the adjusted visible albedo, we convert the visible brightness counts from the satellite, b_i, to visible albedo and divide by the cosine of the solar zenith angle, ϕ, to correct for the time of day:

$$v_i = \left(\frac{b_i}{255}\right)^2 \Big/ \cos(\phi_i). \tag{12}$$

An example of this adjusted visible albedo is shown in Fig. 1. We also remove the constant (over the three months we studied) background albedo that is due to the land surface. This background is calculated as the average of the adjusted visible albedo on clear days in the training set so that

$$z_i = v_i - \bar{v}_i^{clear} \tag{13}$$

The distance metric for the cloudiness correlation parameterization is the absolute value of the difference between pixel values of the adjusted visible albedo image (with the land surface background removed):

$$r_{ij} = |z_i - z_j|. \tag{14}$$

Thus, the cloudiness covariance $\mathbf{P}$ is constructed by applying Eqs. (5)–(7) and (12)–(14) with a tuned k, l, and d for each individual satellite image.

4.5. OI summary

In summary, to perform OI, one must first collect background, $\mathbf{x}_b$, and observation, $\mathbf{y}$, data. Then define observation error covari-