



VIIRS Global Area Coverage (VGAC): A reduced resolution VIIRS dataset for climate data record continuity

Kenneth R. Knapp¹, Philip N. Casey², Yuhua (Douglas) Rao²

¹Knapp WeatherSat Services, Asheville, NC, 28806, USA

²North Carolina Institute for Climate Studies, North Carolina State University, Asheville, NC, 28801, USA

Correspondence to: Kenneth R. Knapp (ken@knappweathersat.com)

Abstract. The Visible Infrared Imaging Radiometer Suite (VIIRS) instrument represents a significant change from the legacy weather satellite observations. For example, there are several differences in scanning strategy, channels, file structure, etc. from the Advanced Very High Resolution Radiometer (AVHRR) which has flown on low earth orbit satellites since the 1970s.

Several longstanding climate data records were developed expressly using the AVHRR instrument's Global Area Coverage (GAC) data. The VIIRS Global Area Coverage (VGAC) dataset is an attempt to more closely bridge the old GAC technology with the new. The VGAC data file provides one orbit of VIIRS observations that have been degraded in resolution to approximate the AVHRR GAC resolution by defining a swath projection model. Each channel is provided using a swath that simulates a scan at ~3.9 km resolution. VGAC provides mean and standard deviation values for each moderate resolution (750 m) VIIRS channel. Higher resolution VIIRS imaging channels (375 m) contain more information, including maximum and minimum radiances, which benefit algorithms like cloud and surface retrievals. The result is that VGAC data are 50 times smaller with 1800 fewer files than the original VIIRS data. The main purpose of VGAC data is to facilitate climate processing to continue the historic climate records from AVHRR and other satellites. VGAC data described here are available from NOAA via <https://doi.org/10.25921/gsef-pg81> (Knapp et al., 2024).

1. Introduction

The meteorological NOAA polar orbiting satellites have provided more than 40 years of observations from the Advanced Very High Resolution Radiometer (AVHRR) instrument. This period of nearly constant instrument observations has facilitated the production of numerous weather and climate datasets over the years. Given the length of the records, the stability of the algorithms, and the vast scientific stewardship applied to the record, many of these products were deemed climate data records (CDRs) (National Research Council, 2004). As such the AVHRR record is a cornerstone of the NOAA Climate Data Record Program (Yang et al., 2016), providing foundational observations for records, that include cloud properties (Heidinger et al., 2014), land surface reflectance (Franch et al., 2017), the polar environment (Key et al., 2016), atmospheric aerosols (Zhao et al., 2008), and more. Furthermore, the EUMETSAT Climate Monitoring Satellite Application Facility (CM SAF) depends on AVHRR for several of their climate products (Hollmann et al., 2006; Karlsson et al., 2023). However, the last AVHRR to be



30 flown was launched on the METOP C satellite, which has an expected end of life of 2027 (World Meteorological Organization, 2026).

More than four decades of observations also means the CDRs incorporate old observing technology. Modern satellites are equipped with instruments that have improved spatial and spectral resolutions. Nevertheless, it is important to continue the AVHRR observational record to sustain the numerous climate observations. While the new instruments have superior
35 environmental data products, the period of record is much shorter. Algorithmic differences between the new products and the historic record can also introduce spurious signals for long-term environmental monitoring. It is important to maintain a consistent historic record using the new data to ensure stability across different satellite instruments.

The VIIRS instrument represents the replacement for the AVHRR instrument (Goldberg et al., 2013). It has 22 channels at a much higher spatial resolution than AVHRR GAC. The increased number of channels and improved spatial
40 resolution lead to a significant increase in file sizes and file numbers. Table 1 compares AVHRR GAC data and VIIRS characteristics. While it is a replacement, VIIRS data are different than AVHRR GAC data in every aspect. Seventeen more channels with resolutions up to 16 times higher result in much more data (an increase by a factor of 400). Even the spatial sampling is vastly different, where VIIRS data is thinned to reduce areas of overlap (this is discussed further in Sect. 2.2). The differences prevent the VIIRS data from being directly used in the CDR algorithms for continuity. The objective of the VIIRS
45 Global Area Coverage (VGAC) dataset is to allow users to directly use VIIRS data with existing AVHRR-based algorithms, thus allowing the continued production of Climate Data Records (Knapp et al., 2019).

This paper describes the VGAC dataset, which acts as this bridge between the legacy long-term AVHRR GAC data records and the newer observing capabilities. The data sources and requirements are described in the following section, followed by the methods used to derive VGAC data. After providing a description of the VGAC data, we summarize recent
50 assessments of the data quality and then provide information on the data availability.

2. Input Data

It is important to understand the legacy GAC observations from AVHRR to provide a context for the newer VIIRS instrument capabilities and the need for VGAC.

2.1 AVHRR GAC Data

55 The AVHRR GAC data derive from 3 separate versions of the same instrument with slight modifications: AVHRR/1 first launched in 1979, but was quickly replaced by AVHRR/2 in 1981. AVHRR/2 remained in operations until an upgrade in 1998 to AVHRR/3. The differences between versions were slight. Some constant aspects include the spatial resolution and a 5-channel instrument (Kidder and Vonder Haar, 1995). Data were scanned at approximately 1km resolution (which represents the full resolution of the AVHRR instrument) and processed on-board the satellite to create, store and transmit a lower
60 resolution for an entire orbit, termed Global Area Coverage (GAC) (Kidwell, 1998; Robel, 2009). AVHRRs 1 and 2 could not



handle entire orbits at full resolution, so they only stored and transmitted smaller, pre-selected regions at the full resolution (called Local Area Coverage). However, they also had the ability to transmit full resolution data in real-time to stations that were in the satellite line-of-sight and had sufficient satellite tracking antennae, software for storage, etc.; this is called High Resolution Picture Transmission. The GAC coverage is the only version of AVHRR with sufficient global coverage for large

65 scale weather and climate applications. Characteristics of GAC data relevant to this discussion are provided in Table 1.

Table 1 - Comparison of key data characteristics between AVHRR GAC, original VIIRS, and VGAC.

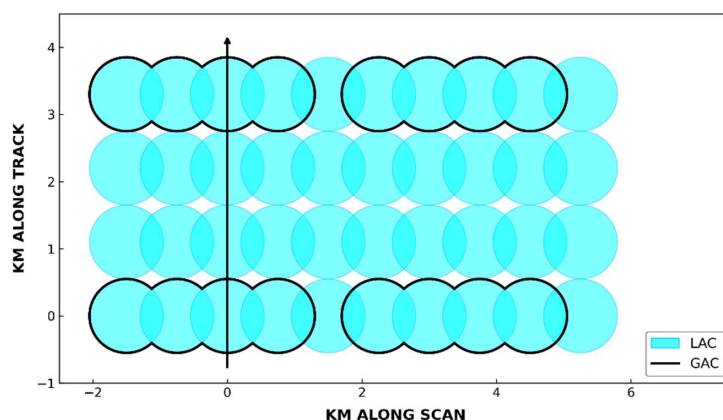
	AVHRR GAC	VIIRS (M band / I band)	VIIRS GAC (VGAC)
Number of files per orbit	1	~1775	1
Number of channels	5	22	22
Number of cross-scan pixels	401	3200 / 6400	801
Number of scan lines per orbit	~12,200	~175,000 / ~350,000	10313
Spatial information	Continuous	Thinned	Continuous
Spatial Resolution	4km nadir (increases toward edge)	750 / 375m (thinned)	3.9km (nearly constant)
Volume per orbit (MB)	~60	~23,500	~460

AVHRR GAC data were provided per orbit with all five channels and metadata in one file. While AVHRR/3 added a sixth channel, it was still only able to transmit 5 channels and for a time alternated between a near IR channel (at 1.6 μm) and an IR (at 3.9 μm) depending on the solar illumination. Geolocation data for GAC files were provided for anchor points, which require spatial interpolation to determine location of each pixel. This requires several steps since interpolation is non-linear at the poles and the dateline, so latitude/longitude points are often re-projected into Cartesian coordinates, interpolated and then projected back to spherical coordinates.

75 The GAC spatial resolution can be considered nominally 4 km. In practice, the onboard processing averages data along the scan line: averaging 4 pixels, skipping a pixel, and then averaging the next 4, and so on. While that processing was being performed two subsequent scans completed. So, the averaging was performed for every third line, which is demonstrated by Figure 1. This complex structure was necessitated by the hardware processing capabilities aboard the satellite. The result is an approximate 4-km resolution product with global coverage. Also, the effective footprint size of GAC increased away from the nadir point due to curvature of the Earth and the cosine of the satellite zenith angle. GAC pixel sampling, geolocation



80 anchor points, and other techniques like efficient bit-packing allowed the GAC data to be provided in efficient orbit files
 (approximately 60 MB per orbit).



85 **Figure 1 – Simulated footprints of the AVHRR full resolution scan points demonstrating capabilities for both Local Area Coverage (blue circles) and Global Area Coverage (black outlines). The arrow demonstrates the sub-satellite path of the instrument. [after EUMETSAT (2011)]**

2.2 VIIRS Data

The VIIRS instrument is detailed by Goldberg et al. (2013) and Zhou et al. (2019). The VIIRS data along with other instruments on Joint Polar Satellite System (JPSS) and Suomi National Polar-orbiting Partnership (SNPP) were reprocessed to improve calibration and other aspects, as described by Zou et al. (2020). The VIIRS represents a significant advance in nearly every
 90 aspect of the legacy AVHRR instrument. The differences discussed here are relevant to the significant limitations of extending AVHRR applications by using VIIRS data. The readers are directed to other sources for a more thorough description of the VIIRS instrument.

The VIIRS instrument provides significantly more data – more channels at a higher spatial resolution – requiring orbits be broken into granules (Table 1). The VIIRS instrument provides information for 22 channels: 16 moderate (MOD)
 95 resolution channels at ~750m, 5 channels at ~375 m resolution (called image channels – IMG) and a day-night band (DNB) with a logarithmic scaled response, providing visible wavelength observations for the entire orbit (at 750 m). Several of the channels have very similar spectral response functions to channels from AVHRR (Hillger et al., 2013). Due to the higher resolution and number of channels, the number of files and data volume per orbit are greatly increased. There are approximately 1800 files per orbit and the volume is almost 400 times larger than GAC, at ~23 GB per orbit. The logistics of processing such
 100 a volume is significant when compared to AVHRR GAC data.

The VIIRS scan structure is detailed by Wolfe et al. (2013). Contrary to the AVHRR (which scans a single line with each sweep), the VIIRS captures 16 MOD detector rows each sweep, often called a push broom technique. The result is that only one row is truly nadir. As this horizontal scan approaches the edge of the swath, the pixel sizes can grow, causing an



overlap between adjacent scans (termed the bow tie effect). To decrease the overlap, data from some of the overlapping portions of a row data are deleted, called data thinning. Another aspect of the data processing is the spatial footprint. An aggregation scheme is used to maintain a relatively uniform pixel size: where the instrument averages 3 samples near nadir for 1 pixel observations, decreasing to two samples, then one sample per pixel near the swath edge. The result is a more uniform pixel size from nadir through the edge of the scan.

The impacts of this observation strategy are significant and result in noticeable spatial discontinuities in the data, as demonstrated in Figure 2. The three boxes each represent relatively the same size: about 33km in width (0.3° latitude). There are three scans of VIIRS shown in 3 colors (blue pluses, magenta squares and yellow crosses); each scan captures 16 rows of MOD channel data. The left image is near the center of the swath with arrows showing the direction toward the swath edge. Meanwhile, the center image is moving East along the swath and shows where the second of the two data aggregation and thinning occurs while the right image shows the eastern swath edge. One noticeable discontinuity is in the nadir image. With a scan rate of 1.78 sec, the Earth rotates 825m in that time near the Equator, which is larger than the MOD resolution of 750 m. Therefore, for a MOD channel, the center pixel (1600) in one swath is adjacent to the next pixel (1601) in the subsequent swath, causing a discontinuity. The thinning creates another discontinuity as the scan continues towards the edge. In the middle image, lines from the full swath (black line) are dropped (going from 14 lines per swath to 12) resulting in a narrower swath (represented by the gray lines) but still with some overlap. The result is missing data in the data arrays (the data arrays from this point on have missing values). Also, the aggregation to maintain a relatively constant pixel spacing results in pixels further apart on western portion (because at the same time when lines are dropped for thinning, the aggregation changes). Lastly, the overlapping nature of the swaths are not fully removed by the thinning. This has its largest impact near the swath edge (right image) where despite the thinning there are still 5 lines overlapping between swaths. Hence, as one moves ‘along the scan’ (in this image, south to north) in the data array, the position moves backward (to the south) when moving between swaths. These discontinuities make spatial processing of raw VIIRS data nearly impossible. The result is a large set of data provided across many files whose structure is not directly usable for processing in algorithms designed for AVHRR GAC data.

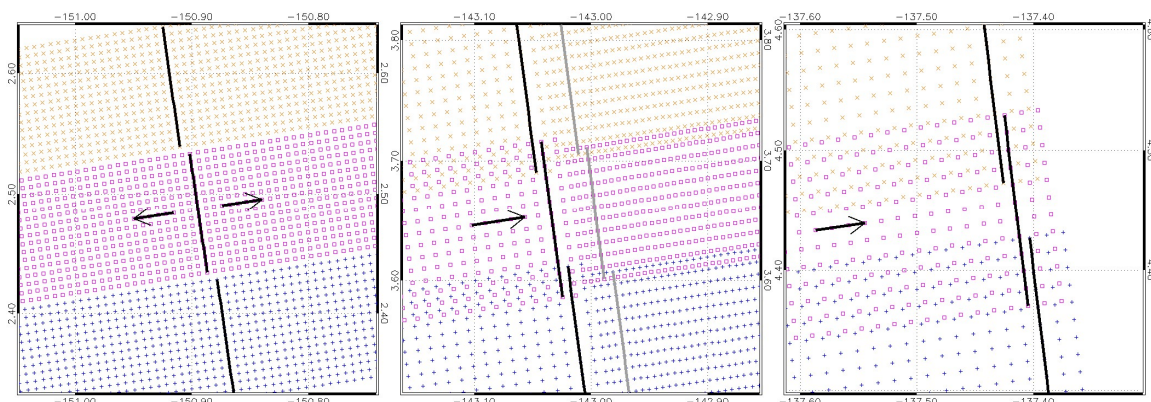


Figure 2 - Three portions of a VIIRS granule near the swath center (left), where thinning starts (center) and at swath edge (right). Each color represents a different scan (blue plus, magenta squares and yellow crosses). The extent of the swath is presented along track by bold black lines. The spatial dimension of each square is the same (approximately 33 km). Arrows point toward the swath edge.

2.3 Requirements for VGAC

The concept of VGAC began with the need to continue the International Satellite Cloud Climatology Project (ISCCP) dataset (Rossow and Schiffer, 1999; Young et al., 2018). The ISCCP cloud tests depend on spatial continuity (Rossow and Garder, 1993) as do other approaches for cloud signals (Coakley and Bretherton, 1982; Saunders and Kriebel, 1988). Hence, the ability to continue those spatial and temporal cloud tests was important. Many changes to VIIRS are needed to continue the historic AVHRR record. The following are some of the requirements needed to continue the ISCCP record:

- **Spatial continuity** - The ISCCP cloud processing assumes that neighboring pixels are spatially contiguous. It uses a region of pixels to determine the variability and the likelihood of a pixel being cloud-filled. VIIRS pixels are not contiguous. The spatial thinning and the spread of the pixels with larger view zenith angles means that some pixels are in opposite directions.
- **Lower spatial resolution** - Original ISCCP processing began with 8 km AVHRR (it also thinned the original data). More recently, processing for ISCCP H uses the full GAC resolution at 4 km. This is much different from the VIIRS IMG channels (~375 m) and the MOD channels (~750m). Changes are needed to make VIIRS data consistent with AVHRR GAC.
- **Channels** - While the ISCCP algorithm only required information from a visible channel and another in the infrared window, all VIIRS channels are provided in VGAC to support other applications.
- **Orbital files** - Data stored by orbit means fewer files and simpler process tracking than thousands of smaller granules.
- **Swath projection** - A new projection is defined that allows the latitude/longitude of a pixel to be explicitly calculated. Likewise, the pixel location in the orbit can be calculated from the Earth location. This is a first for most polar orbiting



data and can accelerate processing (since finding the nearest pixels to a given Earth location can be calculated instead of brute-force searched).

- Complete spatial information - The historical AVHRR GAC data only provided latitude and longitude for anchor points, which increases algorithm complexity (and potential for errors). Storing location information for each VGAC pixel simplifies processing. While not consistent with AVHRR, it is an acceptable and desired improvement.
- Format - The data are reformatted to netCDF-4 with Climate and Forecasting convention to ensure data interoperability and usability.

By providing spatially continuous VIIRS data for all channels, in orbital files, in a portable format and at a lower spatial resolution, we provide a satellite data record that can be used in legacy algorithms, thus maintaining existing satellite-based climate data records.

3. Methods

The primary step for the VGAC process was to define a suitable swath to allow collection of channel information. A well-defined projection allows direct mapping of VIIRS pixels into the VGAC cells. This allows the calculation of statistics for each cell, based on the original channel resolution.

3.1 Defining the swath projection

We chose to have a relatively constant VGAC pixel size. In the along-track direction, we selected 3.9km, which is one third of a VIIRS swath. The VGAC cross-track pixel size was maintained to support a relatively uniform pixel size. The actual width of the VIIRS swath can vary with satellite altitude, which varies from 830 to 856 km (Wolfe et al., 2013). On average, the VIIRS swath is approximately 3000 km, so we defined the size of the VGAC swath as 801 pixels, which corresponds to about 3124 km. Figure 3 demonstrates the VGAC swath projection using the VIIRS pixels first shown in Figure 2. At nadir, the cells clearly show that the three cells match the VIIRS swath (3 cell boxes per swath color). The VGAC cell area is nearly constant across the scan: it varies by only 3% from nadir to swath edge. Compare that to the VIIRS pixel, where the area varies by more than 100% across the swath. Lastly, the edge of the VGAC swath exceeds the VIIRS swath width in some places (resulting in empty boxes in the right plot of Fig. 3) since the actual VIIRS swath is wider when the altitude is higher and narrower when the satellite is closer to the Earth.

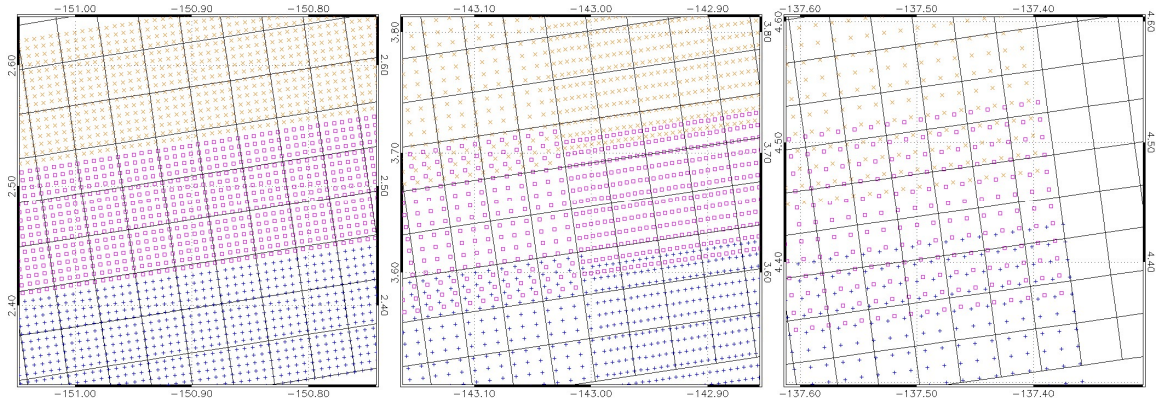


Figure 3 - Same as Figure 2 except the VGAC cell boundaries are now drawn in as black lines.

180 3.2 Remapping VIIRS pixels to VGAC cells

The VGAC processing defines a swath projection in order to remap VIIRS pixels. The VGAC swath is defined by the following variables:

- θ_o , latitude of the swath starting point (at center of orbit swath),
- λ_o , longitude of the swath starting point (at center of orbit swath),
- 185 • φ_o , trajectory of the orbit, in units of degrees east of North,

and the following constants:

- Swath length (number of VGAC cells): $N_L = 10313$,
- Swath width (number of VGAC cells): $N_P = 801$, and
- Cell size: $\beta = 3.9\text{km} / 6378\text{ km} \sim 0.035^\circ$

190 A point in the swath is defined as (i, j) , where i is the location in the scan and j is the number of the row. The corresponding angles in the satellite frame are: $\alpha = (i - 401)\beta$ and $\lambda_s = j\beta$, where α is the angle along the great circle and λ_s is the perpendicular angle away from the great circle. These are converted to Cartesian coordinates using:

$$\tau_s = \begin{bmatrix} \cos \alpha \cos \lambda_s \\ \cos \alpha \sin \lambda_s \\ \sin \alpha \end{bmatrix}$$

The geographic coordinates are obtained with two rotations using θ_o and φ_o :

195
$$\tau_g = \begin{bmatrix} x_g \\ y_g \\ z_g \end{bmatrix} = R_2(\theta_o)R_1(\varphi_o)\tau_s$$

where



$$R_1(\varphi_o) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi_o & \sin \varphi_o \\ 0 & -\sin \varphi_o & \cos \varphi_o \end{bmatrix}$$

and

$$R_2(\theta_o) = \begin{bmatrix} \cos \theta_o & 0 & \sin \theta_o \\ 0 & 1 & 0 \\ -\sin \theta_o & 0 & \cos \theta_o \end{bmatrix}$$

200 Earth latitude (θ) and longitude (λ) are then determined from r_g using:

$$\theta = \tan^{-1} \left(\frac{z_g}{\sqrt{x_g^2 + y_g^2}} \right)$$

and

$$\lambda' = \text{atan2}(y_g, x_g) + \lambda_o$$

where atan2 is related to $\tan^{-1}$ but whose range is $[-180^\circ, 180^\circ]$. Lastly, the Earth's rotation is accounted for by adjusting the

205 longitude using:

$$\lambda = \lambda' - 15t$$

Where t is the scan time in hours since the first scan of the orbit (the result is in degrees since the adjustment factor is $15^\circ/\text{hr}$).

Converting latitude and longitude to VGAC coordinates (i, j) requires the reverse transformation and often needs to be iterated since the scan time is not initially known. However, each VIIRS pixel can be directly mapped to a VGAC cell (i, j) directly

210 using these equations without iteration because the pixel scan time is in the satellite data.

3.3 Channel statistics

Using this mapping technique, data from each VIIRS channel are collected in each VGAC cell to construct the VGAC dataset. Each VIIRS pixel is transformed from its Earth position (θ, λ) to the nearest VGAC pixel (i, j) and statistics are incremented accordingly. Table 2 provides the statistics provided for each variable. The view and illumination angles – view zenith angle, solar zenith angle, view azimuth angle, and solar azimuth angle – are all stored as the mean values for each VGAC cell. Likewise, the mean DNB value is also stored in VGAC cells. The MOD channels record the mean radiance and standard deviation of the radiances from all VIIRS pixels in the VGAC cell. This provides some level of information on the sub-cell variability. Lastly, the IMG channels provide 4 times the number of pixels compared to the MOD channels, so more information – the maximum, minimum, and a sample radiance – is stored for each VGAC cell in addition to the mean and standard deviation. The sampled radiance comes from the VIIRS pixel nearest to the centroid of the VGAC cell. The minimum and maximum radiance values provide particularly helpful information in areas of sparse cloud cover (e.g., open or closed cell cumulus clouds). Importantly, these products (minimum, maximum and sampled radiance) preserve the original observation noise characteristics and native radiometric sensitivity.

Figure 4 shows an example of various VGAC IMG statistics stored for 6 March 2026 over Eastern Florida and the adjacent ocean for one selected visible (IMG-01 at $0.65 \mu\text{m}$) and one selected infrared channel (IMG-05 at $11.45 \mu\text{m}$). Overall,

225

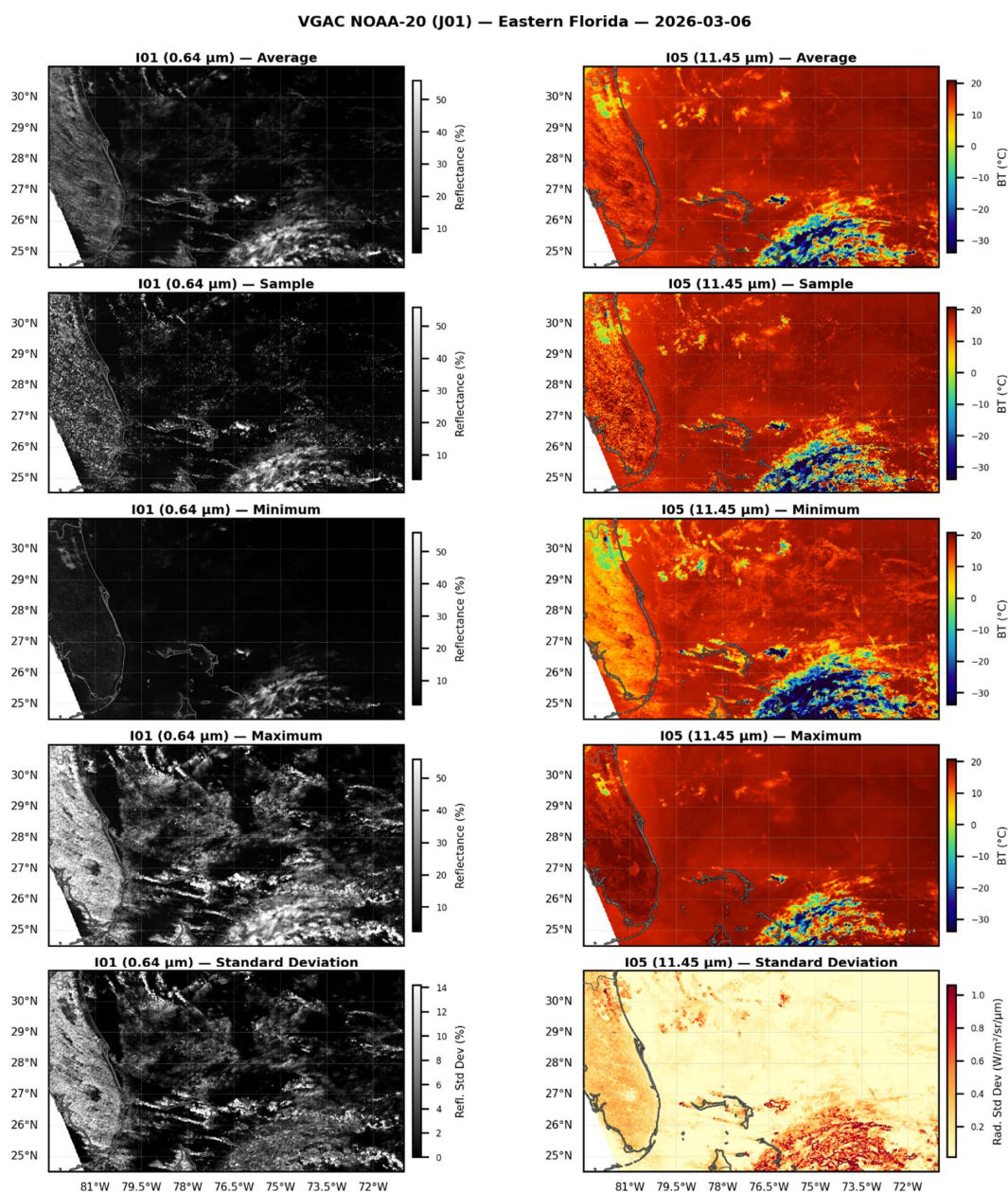


Figure 4 – Demonstration of VGAC statistics provided for two channels (IMG 01 and IMG 05: visible and infrared channels)



there are open-cell cumulus clouds over the land and an area of tropical convection in the southern portion of the image. Over land, the differences between average and sample manifest as blurring of the average image while the sample data appears more speckled, due to differences in how they are constructed. The benefits of the maximum and minimum statistics are also apparent. The minimum reflectance in IMG-01 correspond with the maximum brightness temperatures (BT) of IMG-05. Many of the clouds are smaller clouds because the minimum pixel reflectance corresponds to the surface (e.g., it sees between the clouds). Correspondingly, the maximum reflectance and minimum BT show more clouds. Lastly, the standard deviation of each channel shows how the pixels vary within the VGAC cells.

Table 2 - List of statistics stored in VGAC cells for each geospatial array and different VIIRS channels.

Variable / Channel	Statistics in VGAC Cells
Latitude, Longitude	Mean
View and Illumination Angles	Mean
MOD channels (750m)	Mean, Standard Deviation
IMG channels (375m)	Mean, Standard Deviation, Minimum, Maximum, Nearest to Center
DNB (750 m)	Mean

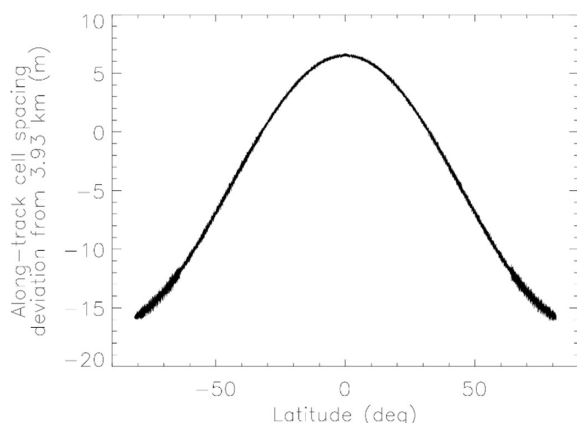
3.4 Assumptions

The following assumptions were made when designing VGAC:

- Constant satellite altitude – The swath projection has no explicit altitude. Yet because it is based on angular constants, it simulates a constant altitude (constant swath width). This results in some missing data at the swath edge when the actual VIIRS swath is narrow (when the VIIRS satellite is at a lower altitude above the Earth). This has no impact on the reported satellite zenith angle because actual values are used to derive the VGAC values.
- Spherical coordinates – Because the VGAC swath is simulated, we used spherical coordinates for simplicity as the angular size of the cells. The constant angle results in a small change in pixel size as a function of latitude (Figure 5). The distance between VGAC “scan” lines varies by only 20 meters, with smaller distances at the poles because the Earth’s radius is less there. Also, mean latitude and longitude values are stored, which leads to the variance (width) of the line in Fig. 5, but only varies by about a meter between “scans”.
- Linear scan time interpolation – As will be noted in Sect. 4, the scan time for each VGAC cell is interpolated. Thus, the differences in reported time and the actual scan time is only 1.2 seconds.



- 250 • Pixel-to-cell assignment – Each VIIRS pixel is assigned fully to one VGAC cell and pixels on cell edges are not split between cells. This has very limited impact.
- Sub-cell statistical variability – Any algorithms requiring sub-cell radiance can use the standard deviation and information from the IMG channels statistics (minimum, maximum and sample radiances) for sub-cell variability.
- 255 • Calibration – The reprocessed VIIRS radiance data (presently: 2012 – 2022) and operational VIIRS calibration (2023 and onward) are provided as-is and without modification which may have implications for downstream algorithm for climate data records.



260 **Figure 5 – Inter-scan line distance for the nadir VGAC cell vs. latitude, presented as the distance (in meters) from a baseline size of 3.93 km.**

3.5 Limitations of VGAC data

Given the assumptions, it should be clear that the VGAC data will not be suitable for every application. To be clear, the initial goal of VGAC is to simulate the historic AVHRR GAC record in order to facilitate continued historic records for climate monitoring. However, it does not thoroughly simulate the GAC record. In most cases, the design was to allow for VGAC to be modified to appear more like GAC by the user but not perform irreversible adjustments. Some significant differences of VGAC from GAC are:

- GAC pixel geometry (i.e., the combined averaging and sampling) is not very representative (as shown in Fig. 1). VGAC is meant to be spatially continuous (as opposed to the in contiguous VIIRS pixels). There was no requirement to fully simulate the GAC-specific 4km spatial averaging and 3 km sampling of GAC 1 km data.
- 270 • GAC pixel sizes grow with view zenith angle and VGAC does not. The VIIRS observation strategy keeps the VIIRS pixels at approximately the same footprint size across the swath. VGAC cell size was defined as an angle ($\sim 0.035^\circ$) providing ~ 3.9 km footprint whose area changes by only 3% between nadir and swath edge. However, if desired, it is possible to spatially average VGAC cells in a manner that simulates the growth of the GAC pixel size. Figure 6 is



calculated by matching the GAC spatial resolution to VGAC spacing. This maps the 801 VGAC cells to the 409 GAC pixels. VGAC and GAC pixels are roughly equivalent near nadir (the variation between 1 and 2 is due to integer round off) and increases to 6 VGAC cells for the GAC pixel at swath edge.

- VGAC radiance data is not adapted to AVHRR spectral response functions. Thus, there will be some level of difference between observations at certain channels between GAC and VGAC. Spectral Band Adjustment Factors (SBAFs) can be applied to specific VIIRS channels to make them similar to AVHRR channels. However, SBAFs are often scene dependent, meaning they are application-dependent. Therefore, this adjustment is left to the user.
- Lastly, there will be certain applications where the VGAC can't be used in place of the original VIIRS data. For example, algorithms or studies requiring high spatial resolution would not be able to use VGAC.

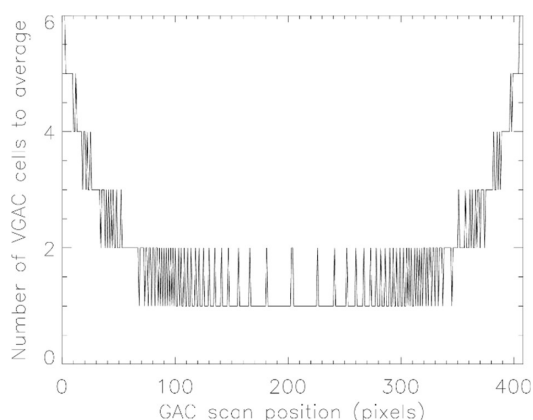


Figure 6 – Number of VGAC pixels to average along the scan line to produce a swath that approximates the GAC pixel growth near the swath edge.

4. Dataset Description

The data for each orbit is stored in a single netCDF-4 file. Each orbit array is 801 pixels wide by 10313 pixels long. Granule files, orbit projection, and channel information are also stored as specific variables in the orbital file.

There are potentially thousands of granules contributing to data for each orbit. For completeness, the file names of granules used for each orbit are recorded as variables in the VGAC file. For example, the “*mod_files*” variable stores all MOD channel granule names used for a specific orbit. Likewise, IMG and ANG data filenames are also recorded to document what granule files were used to produce the VGAC orbit.

As defined in Sect. 3.2, the orbit projection is defined by 4 values: projection time (*proj_time0*), the projection latitude (*proj_lat0*), longitude (*proj_lon0*), and rotation angle (*proj_rot*). This defines the start time of the orbit, the start location (in latitude/longitude coordinates) and the direction of travel at that location. The time of each VGAC scan is interpolated between the start and end time of the orbit. This approximates the 1.78 sec per scan, but because there are 3 VGAC scans for 1 VIIRS



scan, this creates the slightest difference in time: actual scan times should be [0, 0, 0, 1.78] for the first 4 VGAC lines, but instead they are [0, 0.590, 1.18, 1.78]. Using the scan time and the projection, the positions of each grid cell can be directly calculated. However, for simplicity, the center points of each VGAC cell are provided in latitude/longitude coordinates as determined from the mean position of the VIIRS pixels in the cell.

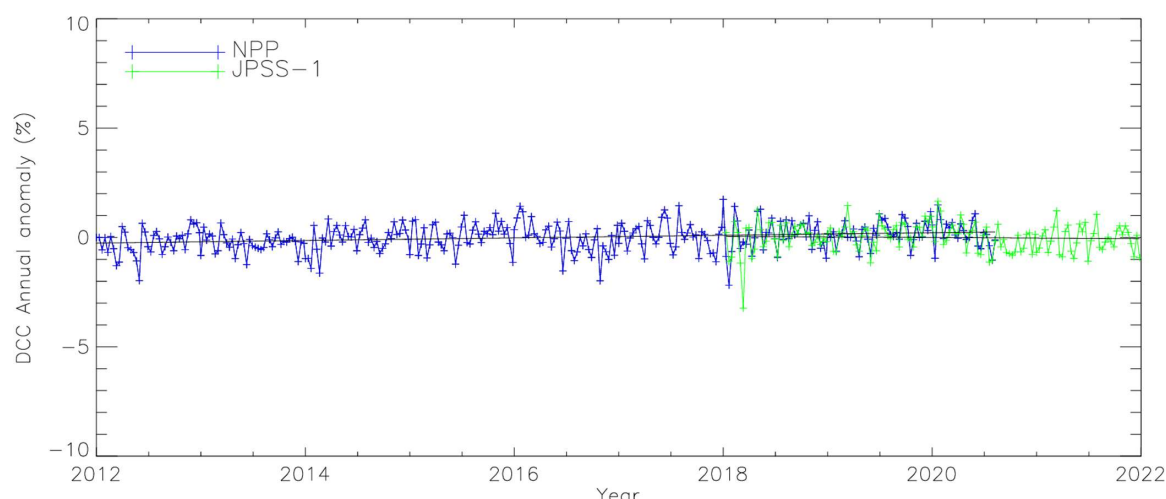
The statistics for the spatial variables (viewing geometry, channel data, etc) are stored in individual netCDF variables which are detailed in Table 2.

Consistent with the source granule data, the radiance variables are stored in scaled integers by converting original float values using netCDF CF convention attributes of *scale_factor* and *add_offset*. For emissive channels, information for the brightness temperature is stored in an array whose indices match the scaled integers stored. Hence the brightness temperature for channel I04 is stored as *i04_LUT[n_lut]* where *n_lut* is 12,000 (which is close to the bit depth of the radiance data). So, the brightness temperature at point (*i,j*) is *i04_LUT[i04_avg[i,j]]*, which means fewer calculations to get BT from radiance.

Despite the added variables to capture necessary information for users, each VGAC orbit file is relatively small. For example, one full resolution orbit of VIIRS is approximately 1800 files and ~23 GB while 1 orbit of VGAC is 450 MB – a factor of 51 smaller. For perspective, that means 1 day of VIIRS granule data has the same volume of 51 days of VGAC. The difference in the number of files is astounding, with a ratio of about 1800 to 1: processing 1 year of a satellite means processing 5250 VGAC files instead of 9.5 million VIIRS files. This significantly simplifies data access and processing for extending legacy climate data record algorithms using the described VGAC data.

5. Data Quality Assessment

The data quality of the VGAC data primarily rely on the quality of the underlying VIIRS Sensor Data Records (SDRs). Wolfe et al. (2013) analyzes VIIRS SDR data from SNPP analyzing geometric and geolocation accuracy finding that (after correction), the pointing accuracy is much lower than the VGAC spatial resolution. Likewise Cao et al. (2013) analyzed SNPP VIIRS calibration, finding visible calibration uncertainty smaller than 2% and IR calibration within 0.1K when compared to MODIS. Both analyzed SNPP data in the first few years of operation. To that end, the data prior to 2023 used the reprocessed VIIRS SDRs data by NOAA Center for Satellite Applications and Research and were denoted as “Final” data. Cao et al. (2021) analyzed nearly 6 years of SNPP radiance calibration data. For visible reflectance channels, they validated long-term calibration using Deep Convective Clouds (DCCs), effectively removing long term trends in reflectance: many channels had trends above 0.1% per year but the recalibration significantly decreased that (after recalibration, only 1 channel still had a DCC trend above 0.1% per year). Likewise, reprocessing improved the stability of the thermal emissive channels: uncertainty decreased from 0.1K to about 0.05K and long term trends in performance are nearly zero (on the order of -0.003 K per year). Thus, we provide some limited analyses below of VGAC as an example of what can be done using VGAC. Also, recent climate analysis with VGAC is summarized as an independent assessment of VGAC data.



330 **Figure 7 - Time series of DCC reflectance from NPP and JPSS-1 for MOD-01, after having removed the annual cycle.**

5.1 Sample DCC processing

The Deep Convective Clouds (DCC) provide a calibration reference for visible channels (Sohn et al., 2009). In general, the statistics for pixels whose top-of-atmosphere (TOA) brightness temperature (T_b) is lower than a certain value are collected and processed. In this case, we collect all pixels where $T_b < 205$ K for DCC analysis (Rossow and Ferrier, 2015) which represents the coldest portions of the clouds. Statistics of the visible information for these DCC cells are collected to assess the calibration. We collect the mean DCC reflectance in 10-day bins and then removed the annual cycle for the SNPP (from 2012 - 2021) and JPSS-1 (2018-2022) periods. This processing benefited from VGAC's reduced number of files and smaller data volume. Figure 7 shows the time series of the entire VGAC record with separated linear trend analysis by satellite platforms. For the overlapping time period between SNPP and JPSS-1, the VGAC time series demonstrates inter-satellite consistency. There is a slight positive trend in the 10-year record of SNPP and an insignificant trend in the shorter JPSS-1 record. But in short, it demonstrates the long-term stability of the MOD-1 reflectance as processed from VGAC. Cao et al. (2021) more thoroughly analyze DCC of all solar reflectance channels.

5.2 Continuity from Polar data processing

VGAC data have already been used to extend and maintain NOAA Climate Data Records and other long term environmental data records. Seo et al (2025) analyzed the impact of using VGAC to extend the PATMOS-x Cloud Properties CDR (Foster et al., 2021). In their analysis, the VGAC radiances show close agreement for 4 of the 5 AVHRR channels, not requiring any adjustment for their processing. Only one (AVHRR channel 2, $0.86 \mu\text{m}$) required significant SBAF. The result on the cloud retrieval algorithms was minor. Some of the resulting differences in water cloud fraction is attributed to VGAC having a higher



spatial resolution than GAC (the cell size doesn't decrease as the zenith angle increases). For cloud optical properties, the differences between GAC and VGAC were not noticeable.

Wang et al. (2025) analyzed the use of VGAC to extend the Polar Pathfinder CDR (Key et al., 2019). They use GAC data to derive various parameters in the polar environment (poleward of 60 deg), producing two sets of data: a set of intercalibrated observations simulating the GAC channels (VIIRS Polar Pathfinder, VPP) and retrievals of geophysical parameters from the VPP (VPP extended, VPP-x) and compare to that derived from AVHRR GAC (likewise called, APP and APP-x). They applied an empirical correction to simulate AVHRR GAC observation, essentially deriving their own SBAF. The resulting visible reflectance differences (between APP and VPP) were ~3.5% and emissive differences ~0.5-0.7K, much of which are attributed to differences in the time of observations (due to orbital drift of NOAA 19). The resulting differences on the derived parameters are small, such that all "biases are within the range of product retrieval uncertainty" (Wang et al., 2025).

Finally, Karlsson et al. (2025) report on the use of VGAC to extend the EUMETSAT Climate monitoring Satellite Application Facility CLARA-A3 cloud algorithm (Karlsson et al., 2023). They analyzed the inter-calibration by deriving their own SBAF to derive simulated GAC channels from VGAC channels using a neural network approach. The resulting retrievals of cloud properties (e.g., optical thickness, effective radius, water paths, etc.) were consistent with original AVHRR-derived counterparts. In fact, they note improved performance of the cloud effective radius retrievals when using a neural network to perform SBAF (vs. the usual linear regression approaches) to adjust VGAC to AVHRR. This is a clear example of why VGAC data is provided without any adjustments: new advances in understanding can improve performance of downstream algorithms. Thus, VGAC is already in wide use and will grow to be used by other climate data records to extend the period of record of other satellite-based datasets (e.g., Knapp et al., 2011; Young et al., 2018).

6. Data Availability

The VGAC data are archived at NOAA National Centers for Environmental Information with digital object identifier: doi:10.25921/gsef-pg81 (Knapp et al., 2024). Additionally, VGAC data is available in the NOAA Open Data Dissemination (NODD) at <https://noaa-vgac-pds.s3.amazonaws.com/index.html>, which provides data freely via cloud object storage with latencies less than 1 day to enable more efficient model data processing workflows. VGAC for NPP has a fixed period of record spanning 2012 to 2021. The records for JPSS-1 and -2 are growing as the operational satellite mission continues, with JPSS-1 (NOAA-20) starting in 2018 and JPSS-2 (NOAA-21) beginning in 2023. Data from follow-on satellites (JPSS-3 and 4) will be included after they are launched and the data become operational.

Data files follow the naming convention:

VGAC_V[sat]02MOD_A[YYYY][JJ][HH][MM]_n[ORBIT]_K005.nc

Where:

- *Sat* can be "N20" for NOAA-20, "N21" for NOAA-21 or "NPP" for Suomi NPP data.



- *YYYY* is the year at the beginning of the orbit
- *JJJ* is the day of year at the beginning of the orbit
- *HH* and *MM* are the hour and minute at the beginning of the orbit (in UTC), and
- *ORBIT* is the satellite orbit number

385 For example, the file “VGAC_VN2002MOD_A2025305_2131_n41223_K005.nc” is from the 41,223rd orbit of NOAA-20, which began on November 1, 2025 (day 305) at 21:31 UTC. Files are stored in netCDF version 4.

7. Code Availability

Production source code for VGAC data is available upon request. The VGAC data production was originally developed in IDL and BASH, was ported to Matlab for wide-scale processing (as implemented by STAR) and most recently was refactored
 390 to run in the AWS infrastructure (lambda processing and s3 storage) using Python for routine processing. Python Jupyter notebook tutorials demonstrating VGAC access, simple data processing, and the map projection are available via Zenodo: <https://doi.org/10.5281/zenodo.19952585> (Knapp, 2026).

8. Conclusions

The VIIRS Global Area Coverage (VGAC) data provide a means to support continued production of existing climate
 395 data records. Deriving requirements from the legacy AVHRR instrument, the VGAC data provide enough similarity with its predecessor to support the extension of existing CDRs at NOAA and EUMETSAT CM SAF using new satellite instruments. The framework is also extendible to other instruments onboarding new polar-orbiting satellites, such as METImage (a Visible and Infrared Imager) onboard the next generation of European Meteorological polar orbiting satellites. The concepts of VGAC (swath projection, storing statistics, etc.) can be used to provide a VGAC-like product for METImage to support legacy satellite
 400 product algorithms for extended data records. This new swath projection could also be extended to applications from other instruments on polar orbiting satellites.

In addition to legacy algorithms, the ability to efficiently reprocess at scale using VGAC should not be overlooked. Cloud compute costs scale with the volume of data processed. Thus, a 50x savings in data volume means a significant reduction in data reprocessing cost. Rather than processing and tracking up to 9 million file reads and access for a year of data, we can
 405 now read and process less than 6000 files for the same year. This scale of processing reduces the entry level for large scale computations for global VIIRS data, enabling operations at research budget scales.

The data are produced for existing satellites and future JPSS satellites can easily be incorporated. Data are provided in a novel swath projection that allows direct conversion between Earth coordinates and pixel location. VGAC data are openly available and archived.



410 **Author contributions**

KK conceived and designed the VGAC dataset, developed the swath projection algorithm, implemented all versions of the processing code, performed the analyses, and wrote the manuscript. YR provided project oversight, secured funding for operational production and manuscript review. PC developed the automated operational processing system, contributed to data production and archival and manuscript review.

415 **Competing Interests**

KK operates Knapp WeatherSat Services, an independent consultancy, and performed this work under contract with NC State. YR is a topic editor for the ESSD journal.

Acknowledgments

The authors thank Changyong Cao (NOAA STAR) and Xi Shao (University of Maryland) for porting the original VGAC code to Matlab and producing the reprocessed VGAC record (2012–2022). We also thank Tom Jackson at EUMETSAT for his independent analysis of VGAC processing in support of EUMETSAT development. Brooke Adams and Hilawe Semunegus (NOAA NCEI) contributed to early efforts toward operational VGAC production. This work was supported by NOAA through the Cooperative Institute for Satellite Earth System Studies under Cooperative Agreement NA24NESX432C0001T101 and NC State Contract. AI tools were used for code refactoring and editorial review during manuscript preparation.

425 **References**

- Cao, C., Xiong, J., Blonski, S., Liu, Q., Uprety, S., Shao, X., Bai, Y., and Weng, F.: Suomi NPP VIIRS sensor data record verification, validation, and long-term performance monitoring, *Journal of Geophysical Research: Atmospheres*, 118, 11,664–11,678, <https://doi.org/10.1002/2013JD020418>, 2013.
- Cao, C., Zhang, B., Shao, X., Wang, W., Uprety, S., Choi, T., Blonski, S., Gu, Y., Bai, Y., Lin, L., and Kalluri, S.: Mission-Long Recalibrated Science Quality Suomi NPP VIIRS Radiometric Dataset Using Advanced Algorithms for Time Series Studies, *Remote Sensing*, 13, 1075, <https://doi.org/10.3390/rs13061075>, 2021.
- Coakley, J. A. and Bretherton, F. P.: Cloud cover from high-resolution scanner data: Detecting and allowing for partially filled fields of view, *Journal of Geophysical Research: Oceans*, 87, 4917–4932, <https://doi.org/10.1029/JC087iC07p04917>, 1982.
- EUMETSAT: AVHRR Level 1b Product Guide, EUMETSAT, Darmstadt, Germany, 98 pp, 2011.
- 435 Foster, M. J., Phillips, C., Heidinger, A. K., and NOAA CDR Program: NOAA Climate Data Record (CDR) of Advanced Very High Resolution Radiometer (AVHRR) and High-resolution Infra-Red Sounder (HIRS) Reflectance, Brightness Temperature, and Cloud Products from Pathfinder Atmospheres - Extended (PATMOS-x), Version 6.0, NOAA National Centers for Environmental Information [dataset], <https://doi.org/10.7289/V5X9287S>, 2021.



- Franch, B., Vermote, E. F., Roger, J.-C., Murphy, E., Becker-Reshef, I., Justice, C., Claverie, M., Nagol, J., Csiszar, I., Meyer, D., Baret, F., Masuoka, E., Wolfe, R., and Devadiga, S.: A 30+ Year AVHRR Land Surface Reflectance Climate Data Record and Its Application to Wheat Yield Monitoring, *Remote Sensing*, 9, 296, <https://doi.org/10.3390/rs9030296>, 2017.
- Goldberg, M. D., Kilcoyne, H., Cikanek, H., and Mehta, A.: Joint Polar Satellite System: The United States next generation civilian polar-orbiting environmental satellite system, *Journal of Geophysical Research: Atmospheres*, 118, 13,463–13,475, <https://doi.org/10.1002/2013JD020389>, 2013.
- 445 Heidinger, A. K., Foster, M. J., Walther, A., and Zhao, X. (Tom): The Pathfinder Atmospheres–Extended AVHRR Climate Dataset, *Bulletin of the American Meteorological Society*, 95, 909–922, <https://doi.org/10.1175/BAMS-D-12-00246.1>, 2014.
- Hillger, D., Kopp, T., Lee, T., Lindsey, D., Seaman, C., Miller, S., Solbrig, J., Kidder, S., Bachmeier, S., Jasmin, T., and Rink, T.: First-Light Imagery from Suomi NPP VIIRS, *Bulletin of the American Meteorological Society*, 94, 1019–1029, <https://doi.org/10.1175/BAMS-D-12-00097.1>, 2013.
- 450 Hollmann, R., Mueller, R. W., and Gratzki, A.: CM-SAF surface radiation budget: First results with AVHRR data, *Advances in Space Research*, 37, 2166–2171, <https://doi.org/10.1016/j.asr.2005.10.044>, 2006.
- Karlsson, K.-G., Stengel, M., Meirink, J. F., Riihelä, A., Trentmann, J., Akkermans, T., Stein, D., Devasthale, A., Eliasson, S., Johansson, E., Håkansson, N., Solodovnik, I., Benas, N., Clerbaux, N., Selbach, N., Schröder, M., and Hollmann, R.: CLARA-A3: The third edition of the AVHRR-based CM SAF climate data record on clouds, radiation and surface albedo covering the period 1979 to 2023, *Earth System Science Data*, 15, 4901–4926, <https://doi.org/10.5194/essd-15-4901-2023>, 2023.
- 455 Karlsson, K.-G., Håkansson, N., Eliasson, S., Wolters, E., and Scheirer, R.: Extension of AVHRR-based climate data records: exploring ways to simulate AVHRR radiances from Suomi NPP VIIRS data, *Atmospheric Measurement Techniques*, 18, 3833–3855, <https://doi.org/10.5194/amt-18-3833-2025>, 2025.
- 460 Key, J., Wang, X., Liu, Y., Dworak, R., and Letterly, A.: The AVHRR Polar Pathfinder Climate Data Records, *Remote Sensing*, 8, 167, <https://doi.org/10.3390/rs8030167>, 2016.
- Key, J. R., Liu, Y., Wang, X., and NOAA CDR Program: NOAA Climate Data Record (CDR) of AVHRR Polar Pathfinder (APP) Cryosphere, Version 2 (2), <https://doi.org/10.25921/X2X1-JR34>, 2019.
- Kidder, S. Q. and Vonder Haar, T. H.: *Satellite Meteorology*, Academic Press, New York, 466 pp., 1995.
- 465 Kidwell, K.: NOAA Polar Orbiter Data User’s Guide, 1998.
- Knapp, K. R.: VGAC Tutorial Jupyter Python Notebooks, v1.0.0, Zenodo [code], <https://doi.org/10.5281/zenodo.19952585>, 2026.
- Knapp, K. R., Ansari, S., Bain, C. L., Bourassa, M. A., Dickinson, M. J., Funk, C., Helms, C. N., Hennon, C. C., Holmes, C. D., Huffman, G. J., Kossin, J. P., Lee, H.-T., Loew, A., and Magnusdottir, G.: Globally Gridded Satellite Observations for Climate Studies, *Bulletin of the American Meteorological Society*, 92, 893–907, <https://doi.org/10.1175/2011BAMS3039.1>, 2011.
- 470 Knapp, K. R., Young, A. H., Inamdar, A., and Hankins, W.: ISCCP and VIIRS: Adapting New Instruments to Extend Historical Climate Data Records, AMS - 2019 Joint Satellite Conference, 2019.



- Knapp, K. R., Adams, B., Semunegus, H., Cao, C., and Shao, X.: Visible Infrared Imaging Radiometer Suite (VIIRS) Global Area Coverage (VGAC) (Version 1), <https://doi.org/10.25921/gsef-pg81>, 2024.
- National Research Council: Climate Data Records from Environmental Satellites: Interim Report, The National Academies Press, Washington, DC, 150 pp, <https://doi.org/10.17226/10944>, 2004.
- Robel, J.: NOAA KLM User's Guide with NOAA-N, -N0 Supplement, 2009.
- Rossow, W. B. and Ferrier, J.: Evaluation of Long-Term Calibrations of the AVHRR Visible Radiances, *Journal of Atmospheric and Oceanic Technology*, 32, 744–766, <https://doi.org/10.1175/JTECH-D-14-00134.1>, 2015.
- Rossow, W. B. and Garder, L. C.: Cloud Detection Using Satellite Measurements of Infrared and Visible Radiances for ISCCP, *Journal of Climate*, 6, 2341–2369, [https://doi.org/10.1175/1520-0442\(1993\)006%3C2341:CDUSMO%3E2.0.CO;2](https://doi.org/10.1175/1520-0442(1993)006%3C2341:CDUSMO%3E2.0.CO;2), 1993.
- Rossow, W. B. and Schiffer, R. A.: Advances in Understanding Clouds from ISCCP, *Bulletin of the American Meteorological Society*, 80, 2261–2288, [https://doi.org/10.1175/1520-0477\(1999\)080%3C2261:AIUCFI%3E2.0.CO;2](https://doi.org/10.1175/1520-0477(1999)080%3C2261:AIUCFI%3E2.0.CO;2), 1999.
- Saunders, R. W. and Kriebel, K. T.: An improved method for detecting clear sky and cloudy radiances from AVHRR data, *International Journal of Remote Sensing*, 9, 123–150, <https://doi.org/10.1080/01431168808954841>, 1988.
- Seo, J., Phillips, C., Heidinger, A., Loveless, D., and Foster, M. J.: Future Satellite-Based Cloud Climate Records: Intercomparison and Consistency between GAC and VGAC, *Journal of Applied Meteorology and Climatology*, 64, 439–454, <https://doi.org/10.1175/JAMC-D-24-0132.1>, 2025.
- Sohn, B.-J., Ham, S.-H., and Yang, P.: Possibility of the Visible-Channel Calibration Using Deep Convective Clouds Overshooting the TTL, *Journal of Applied Meteorology and Climatology*, 48, 2271–2283, <https://doi.org/10.1175/2009JAMC2197.1>, 2009.
- Wang, X., Key, J. R., Moeller, S., Dworak, R. J., Shao, X., and Knapp, K. R.: Extending AVHRR Climate Data Records into the VIIRS Era for Polar Climate Research, *Remote Sensing*, 17, 3495, <https://doi.org/10.3390/rs17203495>, 2025.
- Wolfe, R. E., Lin, G., Nishihama, M., Tewari, K. P., Tilton, J. C., and Isaacman, A. R.: Suomi NPP VIIRS prelaunch and on-orbit geometric calibration and characterization, *Journal of Geophysical Research: Atmospheres*, 118, 11,508–11,521, <https://doi.org/10.1002/jgrd.50873>, 2013.
- World Meteorological Organization: Observing Systems Capability Analysis and Review (OSCAR) Tool, World Meteorological Organization, Geneva, Switzerland, <https://space.oscar.wmo.int/>, 2026.
- Yang, W., John, V. O., Zhao, X., Lu, H., and Knapp, K. R.: Satellite Climate Data Records: Development, Applications, and Societal Benefits, *Remote Sensing*, 8, 331, <https://doi.org/10.3390/rs8040331>, 2016.
- Young, A. H., Knapp, K. R., Inamdar, A., Hankins, W., and Rossow, W. B.: The International Satellite Cloud Climatology Project H-Series climate data record product, *Earth System Science Data*, 10, 583–593, <https://doi.org/10.5194/essd-10-583-2018>, 2018.
- Zhao, T. X.-P., Laszlo, I., Guo, W., Heidinger, A., Cao, C., Jelenak, A., Tarpley, D., and Sullivan, J.: Study of long-term trend in aerosol optical thickness observed from operational AVHRR satellite instrument, *Journal of Geophysical Research: Atmospheres*, 113, <https://doi.org/10.1029/2007JD009061>, 2008.



Zhou, L., Divakarla, M., Liu, X., Layns, A., and Goldberg, M.: An Overview of the Science Performances and Calibration/Validation of Joint Polar Satellite System Operational Products, *Remote Sensing*, 11, 698, <https://doi.org/10.3390/rs11060698>, 2019.

Zou, C.-Z., Zhou, L., Lin, L., Sun, N., Chen, Y., Flynn, L. E., Zhang, B., Cao, C., Iturbide-Sanchez, F., Beck, T., Yan, B., Kalluri, S., Bai, Y., Blonski, S., Choi, T., Divakarla, M., Gu, Y., Hao, X., Li, W., Liang, D., Niu, J., Shao, X., Strow, L., Tobin, D. C., Tremblay, D., Uprety, S., Wang, W., Xu, H., Yang, H., and Goldberg, M. D.: The Reprocessed Suomi NPP Satellite Observations, *Remote Sensing*, 12, 2891, <https://doi.org/10.3390/rs12182891>, 2020.

515