



Turning Night Into Day: The Creation and Validation of Synthetic Night-time Visible Imagery Using the Visible Infrared Imaging Radiometer Suite (VIIRS) Day-Night Band (DNB) and Machine Learning

Chandra M. Pasillas,^{a,b} Christian Kummerow,^a Michael Bell,^a Steven D. Miller,^{c,a}

^a*Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado*

^b*Air Force Institute of Technology, Wright Patterson Air Force Base, Ohio*

^c*Cooperative Institute for Research in the Atmosphere Fort Collins, Colorado*

Corresponding author: Chandra M. Pasillas, chandra.pasillas@colostate.edu,
chandra.pasillas@afit.edu

Early Online Release: This preliminary version has been accepted for publication in *Artificial Intelligence for the Earth Systems*, may be fully cited, and has been assigned DOI 10.1175/AIES-D-23-0002.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

ABSTRACT: Meteorological satellite imagery is a critical asset for observing and forecasting weather phenomena. The Joint Polar Satellite System (JPSS) Visible Infrared Imaging Radiometer Suite (VIIRS) Day-Night Band (DNB) sensor collects measurements from moonlight, airglow, and artificial lights. DNB radiances are then manipulated and scaled with a focus on digital display. DNB imagery performance is tied to the lunar cycle, with best performance during the full moon and worst with the new moon. We propose using feed-forward neural networks models to transform brightness temperatures and wavelength differences in the infrared spectrum to a pseudo lunar reflectance value based on lunar reflectance values derived from observed DNB radiances. JPSS NOAA-20 and Suomi National Polar-orbiting Partnership (SNPP) satellite data over the North Pacific Ocean at night for full moon periods from December 2018 - November 2020 were used to design the models. The pseudo lunar reflectance values are quantitatively compared to DNB lunar reflectance, providing the first-ever lunar reflectance baseline metrics. The resulting imagery product, Machine Learning Night-time Visible Imagery (ML-NVI), is qualitatively compared to DNB lunar reflectance and infrared imagery across the lunar cycle. The imagery goal is not only to improve upon the consistency performance of DNB imagery products across the lunar cycle, but ultimately lay the foundation for transitioning the algorithm to geostationary sensors, making global continuous nighttime imagery possible. ML-NVI demonstrates its ability to provide DNB derived imagery with consistent contrast and representation of clouds across the full lunar cycle for night-time cloud detection.

SIGNIFICANCE STATEMENT: This study explores the creation and evaluation of a feed-forward neural network to generate synthetic lunar reflectance values and imagery from VIIRS infrared channels. The model creates lunar reflectance values typical of full moon scenes, enabling quantifiable comparisons for nighttime imagery evaluations. Additionally, it creates imagery that highlights low clouds better than its infrared counterparts. Results indicate the ability to create visually consistent nighttime visible imagery across the full lunar cycle for the improved nighttime detection of low clouds. Wavelengths chosen are available on both polar and geostationary satellites sensors to support the utilization of the algorithm on multiple sensor platforms for improved temporal resolution and greater simultaneous geographic coverage over polar orbiters alone.

1. INTRODUCTION

Since its inception, satellite imagery has proved an invaluable resource for operational weather forecasters and atmospheric science researchers and is critical in areas with limited surface observations such as coastal regions or open oceans. Forecasters obtain timely and actionable global atmospheric data for observing and predicting hazardous weather conditions through satellite imagery. Cloud and moisture patterns of varying size and time scales are particularly beneficial for making predictions or putting model forecasts into context. Detecting clouds, especially low clouds, at night is more challenging than in the day time due to the absence of visible imagery from solar reflectance and because of the low thermal contrast between the clouds and underlying surface in the infrared (IR) spectrum (Ellrod 1995). The Visible Infrared Imaging Radiometer Suite sensor (VIIRS) Day-Night Band (DNB) captures radiance measurements utilizing lunar illumination and ambient light sources however, most DNB based imagery quality is tied to the lunar cycle. Furthermore, due to the nature of polar orbiters, on which the VIIRS DNB resides, significant lags in revisit time depending on the geographic location of the area of interest may decrease its use for some forecasting elements and time scales. More readily accessible nighttime visible imagery products that are consistent across the lunar cycle can improve forecasters situational awareness of clouds at night, especially outside of the polar regions.

This paper proposes to use machine learning methods with IR channels and DNB lunar reflectance values to create pseudo lunar reflectance values and associated nighttime imagery products, the Machine Learning Nighttime Visible Imagery (ML-NVI). Unlike most DNB scaling methods or

previous machine learning methods for nighttime visible imagery creation, it creates a synthetic lunar reflectance value that can be used for quantitative evaluation against lunar reflectance values directly calculated from sensor measured values. The values and associated digital imagery created from ML-NVI are not an exact translation of lunar reflectance as seen in the DNB for all clouds as it also identifies thin cirrus which is not captured in traditional DNB lunar reflectance, but provides a similar product that performs better than its IR counterparts. Unlike in IR where cold cirrus masks lower level features, ML-NVI detects, but presents a transparent appearance of, some upper level clouds that enables greater fidelity of lower clouds, visually accentuating low-cloud features to the user. Created and validated using only JPSS VIIRS channels, model input channels were selected to be transferable to channels on geostationary satellites in order to gain temporal resolution and increase simultaneous spatial coverage. The model is trained, validated, and tested using Miller and Turner (2009) DNB lunar reflectance as the predictand, permitting both qualitative and quantitative night-time assessments.

The remainder of this paper delves into the development, validation, testing, and evaluation of the ML-NVI model. Section 2 provides essential background information on topics critical to understanding the problem set and development of the algorithm. Section 3 highlights data selection and processing. Section 4 focuses on model architecture selection and parameter tuning, resulting in a preferred model scheme. This scheme is then utilized to create three distinct models based on different latitude ranges that are then qualitatively and quantitatively compared with independent data sets for use with VIIRS in section 5. Section 5 also compares performance over different lunar reflectance ranges and over the full lunar cycle for the model of choice. Section 6 provides case studies for fog and tropical cyclone events that occur outside of the initial training, testing, and validation region as well as conclusions.

2. BACKGROUND

a. *Infrared Cloud Detection at Night*

Visible imagery is especially helpful to forecasters as it offers the highest spatial resolution and presents clouds in a way that matches an analyst's intuition as it is similar to how clouds are observed by the human eye (Kidder and Vonder Haar 1995). Visible imagery provides contextual clues on cloud formations in the way of optical depth and texture. This make it one of the "most easily

interpreted type of satellite products (Kidder et al. 2000).” Due to the lack of solar reflectance, with the exception of data from the DNB, there is no readily available nighttime visible imagery and, in its absence, forecasters must rely on IR imagery for nighttime cloud detection. Near-surface clouds have similar brightness temperatures (BTs) at IR wavelengths as the Earth’s surface beneath making it difficult to detect these clouds (Ellrod 1995). Physical temperature inversions make the detection of low clouds even more challenging (Kidder and Vonder Haar 1995; Liou 2002; Chapman and Gasparovic 2022). At the pixel level, multispectral techniques prove to be more useful for cloud detection over single channels as they account for the property differences of clouds at different wavelengths (Ellrod 1995). Differences between wavelengths highlight low-level cloud features best and cloud spectral properties can be capitalized on through the use of brightness temperature differences (BTDs) in addition to BT values when creating new cloud detection products (Calvert and Pavolonis 2010).

Many previous studies have shown the benefits of using BTDs or multispectral imagery for cloud property determination (Prabhakara et al. 1988; Inoue 1989; Strabala et al. 1994; Ellrod 1995; Anthis and Cracknell 1999; Kidder et al. 2000; Bendix 2002; Ellrod and Gultepe 2007; Hillger 2008; Calvert and Pavolonis 2010; Lindsey et al. 2014; Kim and Hong 2019; Miller et al. 2022). Inoue (1989) demonstrated that combining single wavelength temperature thresholds and BTDs leads to better performance than threshold or multispectral differencing techniques alone. Later, Strabala et al. (1994) compared the $8 \mu\text{m} - 11 \mu\text{m}$ BTD and the $11 \mu\text{m} - 12 \mu\text{m}$ BTD and developed a tool to determine if there was cloud present and what phase cloud it was (helping to indicate cloud height and type) based on these two values. Low cloud split window (SW) approaches typically use the $3.9 \mu\text{m}$ and $10.7 \mu\text{m} - 11.2 \mu\text{m}$ channels (Ellrod 1995; Bendix 2002; Ellrod and Gultepe 2007; Calvert and Pavolonis 2010; Miller et al. 2022). This split window technique capitalizes on the difference in spectral response functions in the IR as well as the fact that the $3.9 \mu\text{m}$ is a combination of reflected and emitted radiation. The short-wave albedo or “fog product” of the 1990s GOES satellites used the $10.7 \mu\text{m} - 3.9 \mu\text{m}$ BTD and revolutionized “fog and liquid water cloud detection at night” (Kidder et al. 2000). The $10.7 \mu\text{m} - 3.9 \mu\text{m}$ BTD remains one of the most commonly used tests for low cloud detection at night. Though these method have greatly enhanced cloud detection at night, they do not compare to that of visible imagery and thus the Operational

Line Scanner (OLS) and Day Night Band (DNB) were created to leverage moonlight for feature detection (Kidder and Vonder Haar 1995; Miller et al. 2006).

b. Lunar Illumination

National Aeronautics and Space Administration (NASA) has observed the spectral irradiance and radiance of the moon since 1996 and lunar radiometric models from this data are used to calibrate on-board sensors of satellites (Kieffer and Stone 2005). Miller and Turner (2009) published a dynamic lunar spectral irradiance data set and model which expedited calculations of irradiance making it possible to calculate a lunar reflectance value (between 0-1) for all measured DNB radiances. This enables quantitative studies of the DNB for conducting comparison calculations similarly to that done with data in the visible spectrum. Min et al. (2017) evaluated nighttime reflectance changes from these irradiances and determined that value differences are less than .05% for water clouds between lunar phase angles. This demonstrated that a DNB based nighttime cloud retrieval will have minimal impacts quantitatively due to the lunar phase, though visually it may present as vastly different. Replicating these reflectance values at full moon can provide measurement that can be further exploited in cloud detection algorithms and imagery.

Nighttime visible imagery gained significant popularity after the development of the DNB. The DNB sensor enables radiance collection from $.5 \mu\text{m}$ to $.9 \mu\text{m}$ and can capture more detailed cloud features at night than LWIR (Miller et al. 2013; Min et al. 2017). The DNB sensor is currently only available on polar orbiters, thus only providing limited benefits to phenomena that evolve rapidly outside of the polar regions. It is very sensitive to low levels of light and uses a 3-step gain system to detect radiances over eight orders of magnitude, utilizing lunar illumination and even air glow as its light source for nighttime imagery and full sun for daytime solar illumination (Miller et al. 2013; Liang et al. 2014; Seaman and Miller 2015; Seaman et al. 2015; Zinke 2017; Line et al. 2018). Normalizing this range for visual presentation is often challenging and all DNB radiances must be scaled post-processing to create imagery usable by the human eye (Zinke 2017). Further details on the DNB can be found in the Algorithm Theoretical Basis Document (ATBD) (Line et al. 2018) or the Users Guide (Seaman et al. 2015).

Three DNB scaling algorithms used to view imagery from measured DNB radiances are the Near Constant Contrast (NCC), high and near-constant contrast (HNCC), and erf-dynamic scaling

(EDS) (Liang et al. 2014; Seaman and Miller 2015; Zinke 2017). The NCC creates a pseudo albedo (Liang et al. 2014; Seaman and Miller 2015), the HNCC a normalized radiance value (Zinke 2017), and the EDS a dynamic scaling of the radiances (Seaman and Miller 2015). These products are enhancements to make static images appear to have a uniform illumination regardless of the lunar or solar influences on the viewed scene (Liang et al. 2014; Seaman and Miller 2015; Zinke 2017; Hoese et al. 2023). In addition to differences between the products, imagery from the full moon versus a new moon for the same product will appear vastly different for similar features leading to different interpretations of the same feature as the lunar cycle progresses (Liang et al. 2014; Seaman and Miller 2015; Zinke 2017; Hoese et al. 2023). Furthermore, in all of these methods, users rely on post-processed imagery, which limits them to a qualitative validation with visual inspections or comparisons; and depending on the specific scaling chosen by a user to highlight different features, products like the NCC can appear visually different based on the maximum and minimum values permitted by the user in the scaled image (Seaman and Miller 2015; Zinke 2017; Line et al. 2018).

c. Machine Learning for Nighttime Visible Imagery

A number of techniques have been developed in recent years to create synthetic nighttime “visible” imagery from IR channels (Chirokova et al. 2023, 2018; Kim and Hong 2019; Kim et al. 2019, 2020; Mohandoss et al. 2020; Harder et al. 2020). Many of these seek to exploit relationships seen in the IR wavelengths with daytime visible imagery (Kim and Hong 2019; Kim et al. 2019, 2020; Mohandoss et al. 2020; Harder et al. 2020). Proxy-visible imagery from Cooperative Institute for Research of the Atmosphere (CIRA) uses scaled VIIRS DNB radiances as truth and multiple linear regression to create its imagery. Proxy-visible is used operationally by the National Hurricane Center (NHC) to aid in tropical storm tracking at night (Chirokova et al. 2023, 2018). Qualitative evaluation of these models have been predominantly in comparison to daytime visible images or nighttime IR (Kim and Hong 2019; Kim et al. 2019, 2020; Mohandoss et al. 2020; Harder et al. 2020). Quantitative metrics have been based on pattern evaluation and structural similarities versus on data from observed measurements.

3. DATA

a. Data Set

1) GENERAL

All satellite data used for this research included VIIRS Sensor Data Records (SDRs) and geolocation data for bands M13 - M16 and the DNB band from the JPSS NOAA-20 and Suomi National Polar-orbiting Partnership (SNPP). Lin and Cao (2019) demonstrated that the relative spectral responses (RSR) in the IR spectrum for VIIRS on these satellite are not identical, differing between 0.18 K - 0.06 K. However, this minimal difference in sensor sensitivities is within documented performance limits and thus data from VIIRS off both systems was included when creating the training and validation data set. Additionally, although there are both seasonal and latitudinal variations of the standard atmosphere profile causing differing RSRs in the temporal and spatial region of study, we chose to accept these differences as within reasonable limits for data. Both of these assumptions allow us to create a model product for use in all seasons and with both satellites.

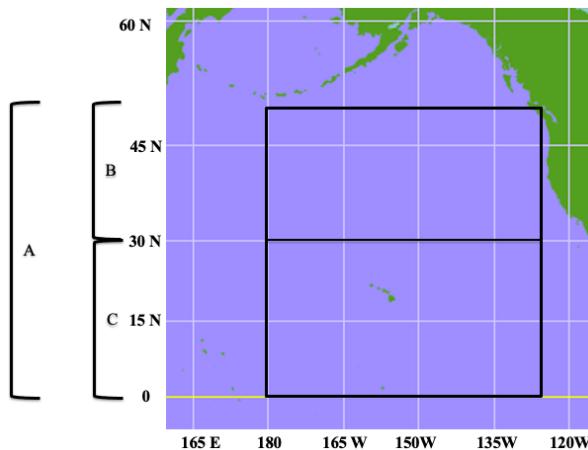


FIG. 1:

Area of Interest (AOI) for the study was a bounding box 50° N, 0° N, 125° W, 180° W [A]. AOI was further divided into two sub-regions [B] and [C] at 30° N. Base map modified from <https://learningweather.psu.edu/node/59>

2) AREA OF INTEREST

Seen in Fig. 1, the area of interest (AOI) used for the training and validation data set as well as the primary testing data set, extends from 0° to 50° N and from 180° W to 127° W. This AOI

was chosen to minimize complications from artificial lights and vegetation differences over land. Equivalent IR channel BTs in different geographic regions (seasons) may be tied to significantly different features: in one region (season) it may represent underlying sea surface temperatures (SSTs) and clear skies but in another region (season) it may represent an area that is obscured by upper-level cirrus or be low lying uniform fog. To account for various background SSTs, which range on average between 5° C to 30° C in this region over the course of the year, and better capture these differences in BTs for similar features, the AOI (0° - 50° N) was further divided into two subsets by latitude range (0° - 30° N, 30° N - 50° N) seen in Fig. 1 as sub-regions C and B respectively.

3) TRAINING AND VALIDATION DATA SET

The training and validation data set was comprised of descending node VIIRS passes over the open northern Pacific Ocean from December 2018 – November 2019 that corresponded to an equatorial pass of 0130 LST on the night of the full moon +/- 1 night.

4) TESTING DATA SETS

The primary testing data set for the final three latitude models covered the same AOI and relation to the full moon as the first data set, but used an independent data set from December 2019 - November 2020. An additional data set from May 21, 2020 – July 20, 2020 was used to qualitatively assess the contrast consistency of ML-NVI imagery across the full lunar cycle. Finally, various additional sets were used for qualitative comparisons for specific weather phenomena to include tropical cyclones and fog. The date-time groups (DTGs) for these sets are provided in the case studies.

b. Data Processing

First, lunar reflectance values were calculated from DNB radiances using the Miller and Turner (2009) algorithm. VIIRS M-band measurements, Miller and Turner (2009) lunar reflectance values, and DNB radiances were then processed with the python package Satpy (Raspaud et al. 2019) to re-project all data to the DNB footprint and corresponding pixels with a nearest neighbor approach. This was done to ensure that the inherent resolution and viewing geometries were consistent. During this process, all data was scrubbed to remove any data sets that had missing values for any

VIIRS BAND	Central Wavelength (μm)/(BTD)	Wavelength Range (μm)
M13	4.05	3.97-4.13
M14	8.55	8.4-8.7
M15	10.76	10.26-11.26
M16	12.01	11.54-12.49
M13M14 BTD	4.05-8.55	*
M13M15 BTD	4.05-10.76	*
M13M16 BTD	4.05-12.01	*
M14M15 BTD	8.55-10.76	*
M14M16 BTD	8.55-12.01	*
M15M16 BTD	10.76-12.01	*

TABLE 1: List of model predictors, the associated central wavelengths or central wavelength brightness temperature differences (BTD) and corresponding sensor wavelength range.

of the predictors or predictand. To reduce any additional bias from the bow-tie effect, training and validation data sets were adjusted to only include data in the first aggregated scan of the DNB, specifically to only include 600 pixels to the left and right of nadir. These pixels represent a total distance of 450 km across the scene. This created 25 sets of 90+ matching DTGs, 12 sets for the training and validation set and 13 sets for the independent testing set. This data could be further divided by month, meteorological season, or latitude if desired.

c. Predictand and Predictor Selection

To best create and evaluate the ML-NVI model, the resulting predictand must be able to be qualitatively and quantitatively compared to a truth. The only nighttime visible data comes from the DNB radiances. While scaling of radiances makes visually appealing imagery to the end user, there is no single standard scaling metric or measurement thus "quantitative assessments on night-time visible imagery are challenging given the qualitative nature of most image products, and details of the scenes used for assessments (Seaman and Miller 2015)". A quantitative assessment of these scaling applications of radiances for digital display would be an assessment of recreating a specific scaling, not against a measured quantity. The Miller and Turner (2009) lunar reflectance is currently the only DNB derived value that can be used for quantitative comparisons for nighttime data from satellite retrievals. It allows for a quantitative evaluation of model generated values used for nighttime imagery that has previously been described only for daytime visible imagery or nighttime IR.

The model predictors and associated wavelengths are seen in Table 1. Initial channel selection was based on wavelengths with similar bands between the VIIRS and ABI sensors. The maximum and minimum values for each wavelength to be used for normalization were taken from the JPSS VIIRS SDR Radiometric Calibration ATBD document Table 2 (Line et al. 2018). The BTDS selected were based on wavelengths relationships highlighted in previous sections. Considerations for channel combinations included all four of Hillger's (2008) recommendations for satellite product development. While experiments were conducted to replicate the current retrieval with fewer predictors, to include not using the BTDS with the understanding that the model would resolve BTD relationships, these models did not produce the same quality imagery as the all-channel all BTD retrieval. All predictors were therefore kept. The final input set for training at the largest AOI data consisted of a 2-D array sized 510,394,368 x 10 representing the number of pixels and number of predictors for data in the prescribed AOI and moon criteria. The independent VIIRS assessment 2-D array was sized 543,686,656 x 10.

4. MODEL ARCHITECTURE AND TRAINING

We trained a Feed Forward Neural Network (FNN) model to create the ML-NVI from IR BT and BTDS using the data set for December 2018 - November 2019. The FNN accounts for non-linearity and processes data pixel by pixel which enables us to further restrict the training data set more easily if needed. As seen in Fig. 2, the final baseline FNN model has ten inputs and is composed of three hidden layers (the first two with eight nodes and third one with four nodes) using Rectified Linear Unit (ReLU) activation and a single node for an output layer that uses sigmoid activation. The model uses an Adam optimizer value of .001 and mean square error (MSE) for its loss function. The number of hidden layers and nodes were chosen after sample runs indicated there was no significant increase to the learning done with further increase in the number of hidden layers or nodes.

Once the baseline model architecture was established, training data was then split 80-20 using the "train_test_split" function from the python model sklearn to accomplish model training (Pedregosa et al. 2011). This baseline model was run repeatedly over various batch sizes and epochs for the three different AOIs. Results of this can be seen in Table 2. It was noted that regardless of batch size and latitude data input, training leveled off after 5-10 epochs and then again around 20-30

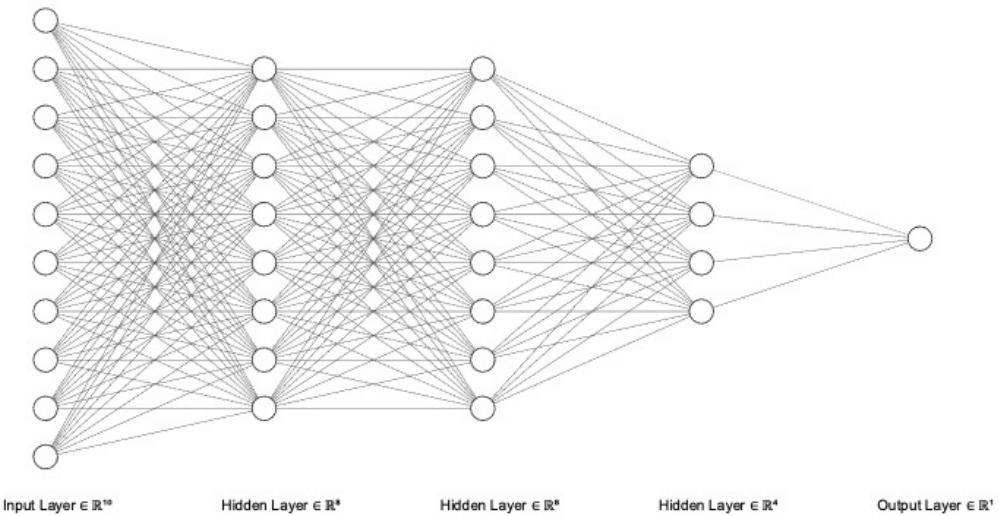


FIG. 2:

Architecture for the ML-NVI. Model Architecture consists of 10 inputs, 2 hidden layers with 8 nodes, 1 hidden later with 4 nodes; all with and Adam optimizer, MSE loss function, and ReLU activation functions; and a single node output layer using the Sigmoid activation function.

epochs. The value gained by these extra epochs was minimal in comparison to the values gained in batch size trade-offs. The metric results between varying the batch sizes between 256, 512, 1024, 2048, and 4096 demonstrated that the value gained from smaller batch sizes was minimal in relation to the additional training time. An epoch size of 40 and batch size of 2048 was chosen for the final model set up. Once the batch and epoch sizes were finalized, training data sets from each latitude range (0° - 30° N (Tropic Model), 30° N - 50° N (Mid-Latitude Model), and 0° - 50° N (Full-Latitude Model) were processed to create a separate latitudinal model for each region (A,B,C). All three latitude models were assessed for their skill over all three latitude regions. The division in training and assessment by latitude enabled an evaluation of the differences in training and model evaluation over tropical versus mid-latitude SSTs, atmospheric profiles, and cloud features. Furthermore, it helped capture instances where equivalent lunar reflectances can be the same but are indicative of different features. This latitude division also aimed to evaluate the benefit, if any, for latitude trained specific models versus a global use model.

Model (epoch/batch size)	Training Latitude	MSE	MAE	V-MSE	V-MAE
10,256	0N-50N	0.0175	0.0872	0.0179	0.0867
	30N-50N	0.0285	0.1212	0.0287	0.1223
	0N-30N	*	*	*	*
30,512	0N-50N	0.0204	0.0988	0.0203	0.098
	30N-50N	0.02828	0.1206	0.028	0.1202
	0N-30N	0.0083	0.065	0.00826	0.0605
30,2048	0N-50N	0.0178	0.0884	0.0177	0.0884
	30N-50N	0.0289	0.1219	0.029	0.1251
	0N-30N	0.0086	0.062	0.0086	0.0625
40, 2048	0N-50N	0.0174	0.0868	0.0175	0.087
	30N-50N	0.294	0.125	0.029	0.125
	0N-30N	0.00847	0.0613	0.00848	0.0622
30, 4096	0N-50N	0.0197	0.094	0.0198	0.094
	30N-50N	*	*	*	*
	0N-30N	*	*	*	*

TABLE 2: The FNN was trained over various epochs and batch sizes for each latitude to determine the preferred combination based on the training and validation mean square error (MSE) and mean absolute error (MAE). The model that had 40 epochs and a batch size of 2048 performed the best overall and most consistently regardless of the latitudes associated with the training data.

5. MODEL EVALUATION AND RESULTS

Loew et al. (2017) stated that “the ultimate goal of a validation exercise is to assess whether a data set is compliant with predefined benchmarks (requirements) that quantify whether a data set is suitable for a particular purpose.” While requirements for the accuracy of radiometers exist and there are desired skill scores for various radiance derived products, there are currently no predefined benchmarks for imagery evaluation. The closest benchmark would be the imagery and validation efforts and analysis methods from NOAA-NASA for the ABI channels. This process first conducts visual inspection for feature determination and temporal image consistency. Next, a qualitative comparison with imagery between legacy and current GOES imagery as well as imagery from polar orbiters is done to assure channels are at least a similar quality. Afterwards, a quantitative comparison of reflectance and brightness values from Level-1B radiances and ground calculated values is done (Pitts and Seybold 2010). Much of the ABI validation efforts are subjective in their feature determination and comparison of quality to other similar products.

The predefined benchmarks for VIIRS imagery are also qualitative and subjective in nature. As highlighted by Hillger et al. (2016), “A major component in the overall strategy for the Imagery

calibration and validation (Cal/Val) effort for the Visible Infrared Imaging Radiometer Suite (VIIRS) is to ensure that the Imagery is of suitable quality for effective operational use. Imagery of sufficient quality is often determined by the ability of human users to easily locate and discriminate atmospheric and ground features of interest.” This demonstrates the highly subjective nature of imagery validation and benchmarks. One user may need a product to highlight upper-level clouds, another need low cloud features, and a third need clear free line of sight to the surface. All may look at the same product and come to different conclusions on its usefulness. Kidder et al. (2000) highlighted that satellite derived products and algorithms need to provide an accurate analysis of the important meteorological features and “do not need to work on the unimportant features of the image.” As an example, they further explained that “an algorithm designed to monitor a low-level feature such as fog/stratus does not need to be precise for high cold clouds.”

The purpose of the study is to create a quantifiable value that represents a physical quantity (pseudo lunar reflectance) and enables resulting imagery that presents clouds at night similar to that of visible or DNB imagery. The greatest focus on cloud detection is for low clouds that are otherwise difficult to detect at night however the product will be qualitatively compared for the cloud types present in the imagery. Quantitative assessments are done for values across the entire scene and ranges to determine overall performance of the synthetic lunar reflectance as a substitute for full moon DNB lunar reflectance values. Independent data from December 2019 - November 2020 was used to evaluate all three latitude models in all three AOIs (A,B,C) resulting in nine sets of statistics and qualitative imagery comparisons.

Similar to Pitts and Seybold (2010), we first conduct qualitative analysis of imagery from the three latitudinal models in comparison to the Miller Turner lunar reflectance, IR, and split window low cloud ($10.6 \mu\text{m} - 4.05 \mu\text{m}$) products to determine preferred model based on visualization of various cloud features. All three models were compared for the entire data set, as well as for specific cases and thus discussion on these may be included for cloud feature assessment. However, inter-model comparisons will be demonstrated for only a few samples with lunar reflectance range comparisons, lunar cycle assessment, and case studies focused on the latitudinal model of choice. Additional imagery for more comparisons between the three latitudinal models is available upon request.

a. Inter-Model Comparisons

1) QUALITATIVE ASSESSMENT

The first comparisons, seen in Fig. 3 and Fig. 4, are done between DNB lunar reflectance, one VIIRS IR M-band, M13 (4.05 μm), the split window (10.76 μm - 4.05 μm) and the three latitude ML-NVI models at the full moon. Fig. 5 also provides a sample of all three latitude models but includes M15 (12.01 μm) instead of the Split Window. The focus for the imagery is for use over open ocean and coastal regions, however the case samples do include some land features and clouds over land. Each latitude model addresses similar cloud types differently because the models were only trained on data (background SSTs and cloud types) specific to that region. The impact of latitude specific trained models and their usefulness for other regions is most apparent in the top left corner of Fig. 4 where the tropical model has created additional cloud cover over the clear ocean due to the SSTs. The contrast is apparent in the overall contrasts between the three models in Fig. 4 - Fig. 6. The added contrast provides the ability to gain more depth of clouds in the ML-NVI over the other IR channels and products. This demonstrates the potential to catch a forecasters' eye more easily when they are reviewing imagery and make a better assessment of the cloud features in question.

Beginning with existing products, Fig.3 highlights differences between the DNB Lunar Reflectance, M13, and the Split Window (10.76 μm - 4.05 μm)in how each addresses the ship tracks that are centered on 39° N and 123 ° W. While present in all imagery except the SWIR, their signal is strongest in the lunar reflectance, the mid latitude ML-NVI, and the full range ML-NVI.

The model trained on data from 0° - 30° N has the least contrast of the three models over these samples. This is because the observed BTs in the predictors are much colder than the tropical SSTs that represented clear skies so the model assumes that there must be interference to the surface and the BT interpretation is a reflectance indicative of a constant thin cloud layer over cold SSTs. This results in lunar reflectance values that are in a smaller range and over forecast reflectance values for clear skies. This model did have decent contrast and performance in the tropical cyclone (TC) examples and scenes over warm SSTs. This demonstrates that this model will only perform well with warm background SSTs. The model trained from 30° N - 50° N appears to have the greatest contrast of the three models and may be the closest to the DNB at first glance. This is because these images cover a majority of the latitudes that this model was trained on.

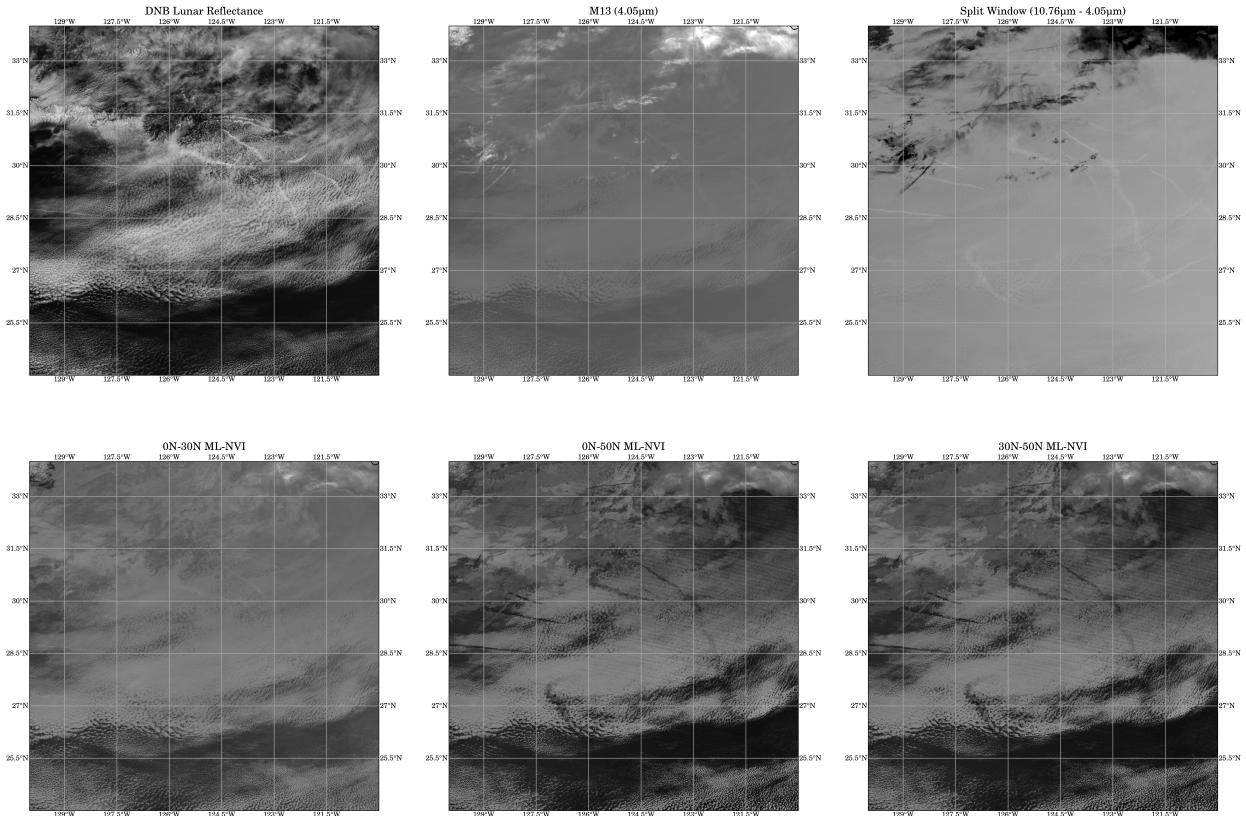


FIG. 3:

Qualitative Comparisons (Left to Right) of images from DNB radiance calculated lunar reflectance, predicted lunar reflectance for the three latitudinal models (Full latitude, Tropical Latitudes, Mid-latitudes), VIIRS M13 (4.05 μm) and VIIRS Split Window (10.76 μm - 12.01 μm) sensors. Image is from the Northern Pacific open ocean on December 12, 2019 at 1026 UTC.

Thin cirrus is absent in the DNB, appears bright white in the IR blocking the underlying clouds, and appears as a semi-transparent mask that you can “see through” to the clouds beneath in the three ML-NVI models. Due to its transparency, we can still detect low to mid clouds in layers below it that are absent or blocked in the IR imagery.

Additionally, all three latitude models capture low level clouds, such as open ocean stratocumulus, that are visible in the DNB lunar reflectance but not captured well, if at all, in the IR samples as seen in the bottom of all images in Fig. 3, with best contrast in the full-latitude model. Though not as bright as DNB lunar reflectance, it is possible to infer the layers of clouds in the models similar to the DNB with higher/thicker clouds being brighter (colder) in comparison to lower and

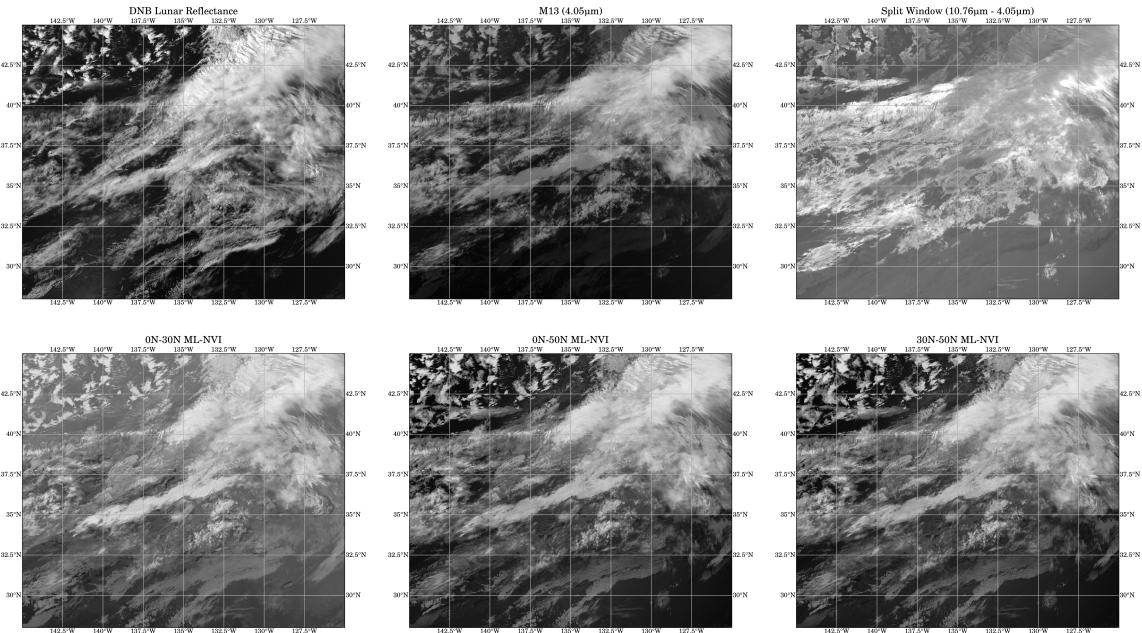


FIG. 4:

Qualitative Comparisons (Left to Right) of images from DNB radiance calculated lunar reflectance, VIIRS M13 ($4.05\mu\text{m}$) and VIIRS Split Window ($10.76\mu\text{m}$ - $4.05\mu\text{m}$) sensors for the top row and predicted lunar reflectance for the three latitudinal models (Full Latitudes, Tropical Latitudes, Mid-latitudes) on the bottom row. Image is from the Northern Pacific of the Western Coast of the United States and over open ocean on December 11, 2019 at 1044 UTC.

with more contrast for lower-level features than seen in the IR. This assessment holds true with the exception of the previously mentioned cirrus clouds that are optically thin and appear translucent but darker than thick cumulus. Through additional imagery assessments it was determined that although the contrast varies between the three latitude models, all models are able to capture ship tracks and provide better texture of clouds to help indicate actual cloud feature versus a uniform field as seen in IR. Figure 5 is used to highlight the differences in model performance over warmer SSTs found in the tropical regions and demonstrates its capabilities and limitations with convective activity. In this case the mid-latitude model captures the in the lower right corner the best however, it struggles with contrast for the lower cloud features in the tropics just like the tropical model did at mid-latitudes.

In comparison to the single channel IR wavelength imagery, the human eye can visualize the low-level cloud features better in the models. As model predictions are created from thermal

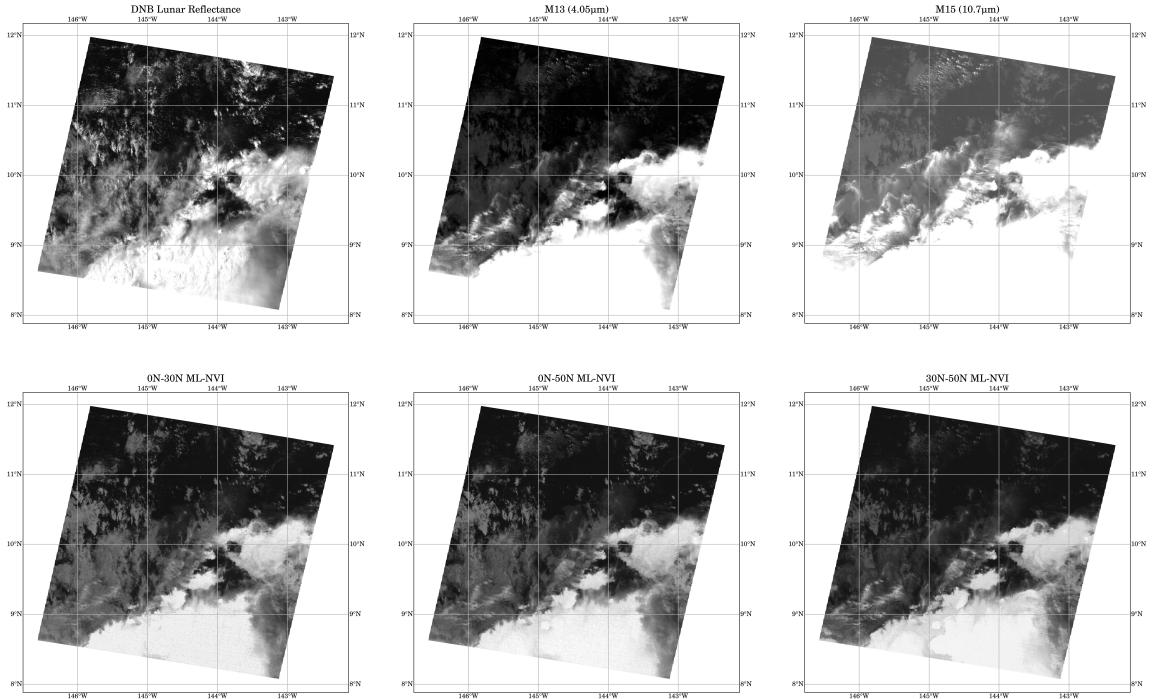


FIG. 5:

Qualitative Comparisons (Left to Right) of images from DNB radiance calculated lunar reflectance, VIIRS M13 ($4.05\mu\text{m}$) and VIIRS Split Window ($10.76\mu\text{m} - 4.05\mu\text{m}$) sensors for the top row and predicted lunar reflectance for the three latitudinal models (Full Latitudes, Tropical Latitudes, Mid-latitudes) on the bottom row. Image is from the Northern Pacific at lower latitudes over open ocean between Mexico and Hawaii for December 12, 2019 at 1113 UTC.

emissions, the contrast between clear skies and clouds do not appear as vividly as the true DNB lunar reflectance. The clouds appear brighter than the clear skies and intuition would show low and layered clouds, missed by IR interpretation alone, are present. An additional example of texture and this contextual clues to forecasters is in the lower right corner of Fig. 4 in which all models capture clouds similar to in the DNB lunar reflectance while the texture is minimum in the IR and split window. The models are also able to remove man-made lights seen in Fig. 10 most likely due to the lights not having a signature in the IR channels as it does in the visible. In the absence of city lights, populated areas appear darker (warmer) than the surrounding land which may be capturing the presence of urban heat islands. More investigation must be done before determining the performance of the ML-NVI for clouds over land.

Qualitatively, the full latitude model performed best for all the cloud types and accounts for all SSTs between 0° - 50° N that are seen in open oceans, appearing similar to the low-latitude model in the tropics and to the mid latitude model in the middle-latitudes.

2) QUANTITATIVE ASSESSMENT

An evaluation of the underlying data distribution for each model is done before evaluating the specific sets of observational points. Comparisons of the distribution of the three latitude model predictions for the validation sets relative to each other and the true reflectance can be seen in Fig. 6. Note that the lunar reflectance value distribution is not normally distributed. The probability density functions (PDF)s further highlight not only the percent of observations types (i.e. reflectance values) but also visually demonstrates differences in cloud detection ability at various reflectance values and by models. A difference in the value density when compared to the DNB distribution is most present in areas of low reflectance and common for all three models. As noted in the qualitative assessment, this shift is partly due to the capturing of cirrus clouds by the ML-NVI. All three models appear to have performed similarly in this shift and throughout the range of reflectances in the PDF, however, visually we see how the specific model differences manifest when turned into imagery. From approximately 0.1 - 0.35 reflectance values, the low latitude model under-forecasted values while the mid-latitude model over-forecasted values and above 0.35 the inverse is true. The middle reflectance values are the hardest to detect and may be indicative of layered clouds, low or mid-level clouds, or scenes where the cloud features are smaller than satellite pixels. The full latitude model appears to perform best in these cases. The highest reflectances represent optically thick clouds and thunderstorm clouds that are typically well identified in the IR due to the strong thermal contrast. A cursory look of the PDFs indicate that the full latitude model has the greatest distribution similarity to the truth across the reflectance spectrum. Important to note in the PDF is the location of the spikes in the larger lunar reflectance values seen across all data sets. It peaks at 1 in the lunar reflectance as this is a maximum set as reflectance, in theory, would not extend beyond this. However, due to scattering, signal noise, and city lights, measurements may exceed 1. Under-forecasting trends and biases can be seen in the models based on the location of this peak and can give insight into the range of values each model may predict.

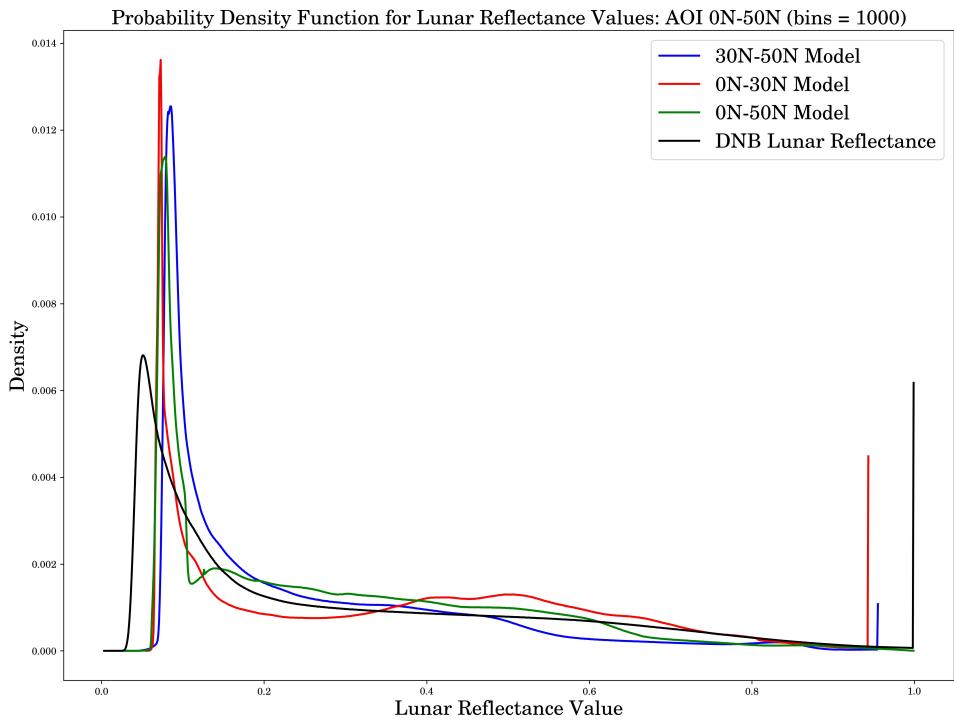


FIG. 6:

Probability Density Functions of the three latitudinal model lunar reflectance distributions versus the true DNB lunar reflectance distribution for the 0° - 50° N validation AOI.

Next, statistical calculations were conducted to assess the overall capability of the ML-NVI for full moon scenes over open ocean for all three models at all three latitude ranges. The product is designed to function regardless of moon phase, however, model creation and quantitative validation was conducted near the full moon, as this is when the highest quality DNB data is available.

During visual inspection of the qualitative imagery it was observed that sections of the lunar reflectance imagery patches were fully black although there was no missing data in the raw DNB radiances. It was determined that this was due to the specific moon angles for which the Miller and Turner (2009) lunar reflectance is not calculated, thus all data-points where the DNB Lunar Reflectance Truth was either 0 or NAN were removed. These points account for approximately 2% of the evaluation data and, while present for imagery creation, were removed before statistical calculations were performed and PDF distributions created.

2020 ML-NVI VALIDATION METRICS									
Model	Latitude Range	Explained Variance	R2	RMSE	MAE	Spearman Corr.	CC	KLD	JSD
0N-50N Full Latitude	0N-50N	0.76	0.76	11.83	8.2	0.86	0.87	0.08	0.14
	30N-50N	0.67	0.67	14.42	11.34	0.82	0.82	0.08	0.15
	0N-30N	0.8	0.79	10.20	6.35	0.87	0.9	0.07	0.13
30N-50N Middle Latitude	0N-50N	0.72	0.682	13.56	9.0	0.85	0.85	0.09	0.15
	30N-50N	0.68	0.657	14.63	10.99	0.83	0.83	0.08	0.14
	0N-30N	0.71	0.655	13.19	7.98	0.85	0.85	0.1	0.15
0N-30N Low Latitude	0N-50N	0.64	0.577	15.68	10.77	0.8	0.81	0.11	0.17
	30N-50N	0.51	0.166	22.85	18.24	0.71	0.72	0.13	0.19
	0N-30N	0.83	0.83	9.33	6.23	0.87	0.91	0.06	0.12

TABLE 3: A consolidation of the validation metrics for all three latitude models across all three latitude validation regions. Scores for the full latitude model in the 0° - 50° N AOI demonstrate model performance for use of a singular global model across all seasons and latitude.

Quantitative scores for the three models are provided for the overall data sets and thus there is unintentional weighting based on the reflectance distribution. A chart of the statistical results for all three latitudinal model assessments at the three validation latitudinal bands is provided in Table 3. There is currently no other published data utilizing lunar reflectance to provide baseline metrics for direct comparisons. As nighttime visible data becomes available, it will be critical to ensure comparison metrics are made between similar data sets. For this reason, metrics are based on all valid pixels and not currently divided by cloud type or height.

The three models were compared using the following metrics: explained variance, R-squared, root mean square error (RMSE), mean absolute error (MAE), Spearman Correlation, correlation coefficient (CC), Kullback-Leibler divergence (KLD), and Jensen Shannon divergence (JSD). As expected, performance metrics were generally best when individual latitude models are evaluated on the same latitude ranges they were trained on. Still there was only a 12% difference in the explained variances between the best and worst models for a latitude. When looking at overall error, lunar reflectance was evaluated from 0-100 and the differences ranged between 3 and 4 for MAE with the exception of the tropical models performance over mid-latitude ranges which was significantly worse than the mid-latitude and full-latitude models for this area. It is important to note that the model performance on the full-latitude data set may also be biased towards the percentage of data points made up by each latitude range.

Using the full set of predictors the model has an explained variance value of 0.76. While this is a good value, especially as it only considers spectral relationships, inclusion of a predictor with information on spatial context may further improve skills in model variance. The ranges of mean absolute error (MAE) and root mean square error (RMSE) further demonstrate the spread of the data and are within a reasonable amount when considering that lunar reflectance is a visual proxy for total cloud cover which is often measured in octas. The Spearman's Correlation and correlation coefficients between the ML-NVI and DNB true lunar reflectance range between 82 and 89% indicating a strong positive relationship between the two. To further quantify this relationship, KLD and JSD scores were calculated. JSD scores range between 0-1 with the lower number indicating the similarity or minimal divergence between the two sets. The JSD scores for the full latitude model ranged between 0.12 and 0.15 and indicate there would be only a small adverse impact if the ML-NVI values were to be used in place of the DNB lunar reflectance for modeling or calculations if using the full latitude model to create lunar reflectance and these are only 0.01 higher than the scores obtained by the sub-latitudinal models as demonstrated in Fig. 7 with the full lunar cycle. This highlights the ability to have a greater visibly intuitive product with a scaled consistency for forecasters across the lunar cycle as the ML-NVI predicts what a full moon lunar reflectance would be regardless of the actual moon phase, angle and existing lunar reflectance and the ability to use one model (full-latitude) for the whole globe. It is possible to infer what lunar reflectances would be expected for the observed phenomena at any lunar phase if a full moon existed instead based on the lunar reflectance values created from our model. From this it could be possible to even utilize daytime cloud mask algorithms at night if the day and night reflectance differences in the 3.9 μm band are accounted for (Miller et al. 2022).

3) MODEL OF CHOICE

Based on qualitative analysis over imagery in the AOI covering all full moon periods from December 2019 - November 2020, and model prediction metrics, the full latitude model (0° - 50° N) is the best performing model overall. For mid-latitudes it holds its own against the 30° N - 50° N (mid-latitude) model and significantly outperforms the 0° N - 30° N (tropical) model. In the tropical regions with warmer SSTs, it is very similar in model performance to the tropical model while the mid-latitude model has the poorest performance. Quantitatively, the full latitude

model reflectance values are the most closely aligned to the DNB reflectance and provided the best representation of clouds among the three models with minimal loss of information. Because of this, the full latitude model is the model of choice, and the remainder of model validation and future case studies will be on the full latitude model, hereby referred to as the ML-NVI.

b. Model of Choice Performance

1) LUNAR REFLECTANCE RANGE ASSESSMENT

Overall model performance was addressed during inter-model comparisons but an assessment of the chosen model's performance over ranges of reflectance values is also useful. In addition to the previously discussed PDF in Fig. 6, a scatter plot of the full latitude model's performance versus the true DNB lunar reflectance can be seen in Fig. 7. The lower left corner represents clear skies with lunar reflection values in truth and predictions of zero and the upper right represents lunar reflectivity values of one in each. The one-to-one line and model best fit lines are also noted. The scatter plot shows a positive trend between the data sets and, though in some lunar reflectance value ranges they appear to be much more widely distributed along the one-to-one line, the relative differences in the data set are small.

Because of data distribution and interest in performance over specific data ranges, we look to assess differences in lunar reflectance values between the truth and model at multiple lunar reflectance ranges. The World Meteorological Organization (WMO) suggests that for cloud cover comparisons, data and models/predictions be divided into cloud categories versus a continuous scale, though they do not specify category thresholds (Zhongming et al. 2012). WMO height and total cloud cover categories for observations and terminal aerodrome forecasts usually line up with aviation flight safety requirements (Weiss 2001). A review of research shows that most studies divide the total cloud cover into observational categories based on sky octa obscuration (clear, few, broken, scattered, and overcast), into 3 categories based on cloud amount (clear, partly cloudy, cloudy), or into 10% bins for precise measurements of cloudiness (Warren et al. 1988; Kidder and Vonder Haar 1995; Hogan et al. 2009; Zhongming et al. 2012). When using satellite, threshold techniques are a common way to delineate between cloud amount categories and transfer well for reflectance values (Kidder and Vonder Haar 1995). Solar reflectance values for clear skies over the open ocean range between 0 - 0.2, thus these values would represent clear skies in lunar reflectance

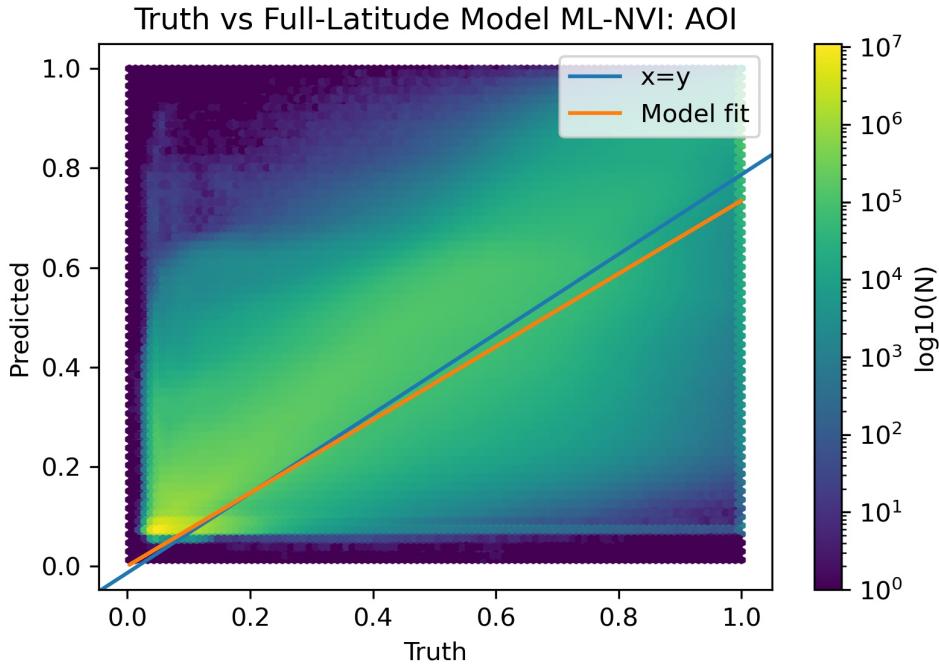


FIG. 7:

Data comparison of the 2020 validation data and the DNB true lunar reflectance for the full latitude model over the 0° - 50° N AOI

as well, while pixels with values above that would have some form of clouds (Ackerman et al. 1998; Kim et al. 2017).

Data was divided into five even ranges of true lunar reflectance values (0-0.2, 0.2-0.4, 0.4-0.6, 0.6-0.8, and 0.8-1) and value differences between the truth observed lunar reflectance values and full latitude model predicted lunar reflectance values as well as the mean and standard deviations are seen Fig. 8 for the five sub-ranges. The percentage of the overall data based on bin ranges are shown in Table 4. Considering this breakout, and referencing Fig. 6, a large amount of the data set included clear skies, thus overall data set standard deviations may be biased towards the models performance on the lower lunar reflectance values.

Dividing the data set into these ranges of lunar reflectance versus looking solely at the overall data set deviations, enables an additional understanding of model performance. Lunar reflectance values less than 20 usually represents clear skies however, they can also represent the presence of optically thin clouds such as cirrus, which the DNB does not detect but the ML-NVI does as seen in the qualitative imagery assessment, or low stratus. As reflectance values increase, optical depth

2020 DATA SET LUNAR REFLECTANCE RANGE VALUE COMPOSITION	
True Lunar Reflectance Value Range	Percentage of Data Set
(0,0.20]	50.5
[0.2,0.4]	19.9
[0.4,0.6]	15.7
[0.6,0.8]	10.1
[0.8,1]	3.8
[0 - 1]	100

TABLE 4: Data set composition across five ranges bins used in evaluating skill for lunar reflectance differences.

values also increase and this could be due to layers clouds or growing cumulus. This evaluation is not a direct assessment of how the model performs for various cloud types or cloud heights but does provide insight into performance for different lunar reflectance ranges in which performance can be inferred (Liou 2002; Kidder and Vonder Haar 1995). In general, the larger the values of lunar reflectance the greater the standard deviation in truth-model values. Additionally, at lower values the model tended to predict greater lunar reflectance values more often than in the DNB lunar reflectance but as the true lunar reflectance values increased it tended to under predict the reflectance values more and with a greater deviation.

2) LUNAR CYCLE ASSESSMENT

In addition to analysis at full moon, a qualitative assessment was done on ML-NVI performance over the full lunar cycle for the model of choice, the full-latitude model. The DNB sensor gathers radiance values from sources other than the moon (airglow, city lights, aurora, etc); because of this some DNB products (HNCC, NCC, ERF scaling) can create images regardless of the moon phase but additional processing must be done and imagery is not standardized to retain the same contrast across the lunar cycle. Lunar reflectance cannot be calculated for all phases as the Miller-Turner calculations require the moon to be at a specific lunar zenith angle. Figure 9 makes comparisons between DNB lunar reflectance imagery, the full-latitude model imagery, and M-band 15 (12.01 μm) for two lunar cycles over the Hawaiian Islands, a region of warmer SSTs. During half of the lunar cycle, lunar reflectance data is not able to be calculated, thus imagery based on reflectance is not available (as seen by the black images for lunar reflectance in Fig. 9 rows 1 and 3) and other DNB scaled products must be used. The ML-NVI provides lunar reflectance values and

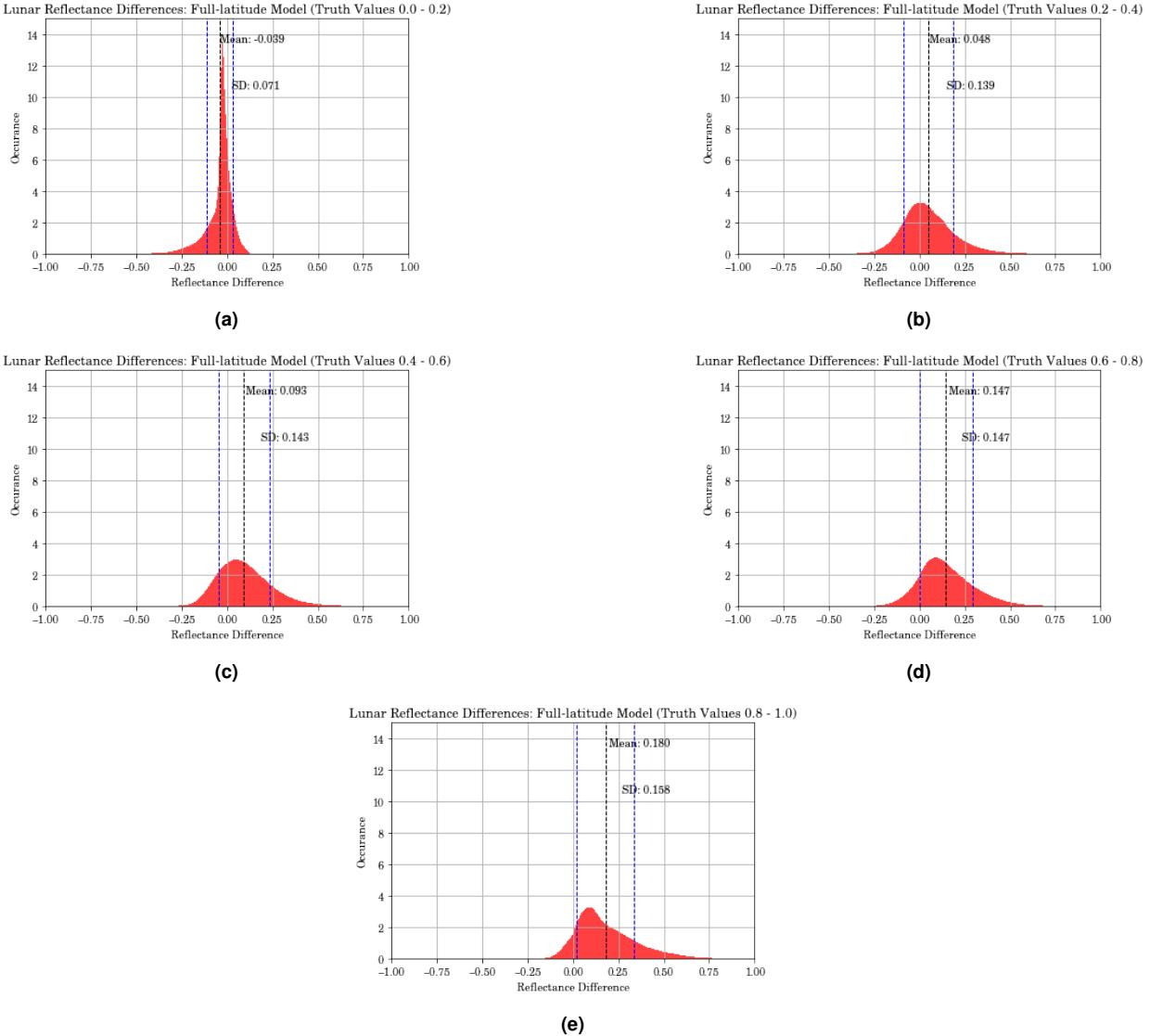


FIG. 8: Distribution of DNB true lunar reflectance - full-latitude model reflectance differences evaluated at 0° - 50° N AOI for truth lunar reflectance values between 0.0-0.2 (a), 0.2-0.4 (b), 0.4-0.6 (c), 0.6-0.8 (d) and 0.8-1.0 (e)

visual imagery similar to that of DNB that would appear at full moon periods but with consistent shading across the lunar cycle as seen in the center column of Fig. 9. Since IR is fully emissive and the model was trained solely on full moon scenes, calculated model reflectance values do not respond to lunar cycles and ML-NVI can provide visually consistent night-time visible imagery, even when DNB lunar reflectance is not available, that enhances cloud identification over IR alone, allowing users to learn only one presentation of features for the full lunar cycle versus varying representations as seen in most DNB enhancements. In comparing the ML-NVI to the IR over the

mountains of the Big Island (located in the region bounded by 19° N, 20° N, 156° W and 155° W), as seen in the second row, the ML-NVI does not interpret these cold land features as clouds with large cold signatures.

6. CASE PRESENTATION AND CONCLUSION

After a general qualitative and quantitative assessment of the ML-NVI's ability to detect cloud and clear skies across the full lunar cycle, ML-NVI was applied to specific meteorological phenomena where the improved detection of low clouds may benefit observers and forecasters. ML-NVI can enhance the detection of low clouds at night over traditional IR, which is especially critical for fog formation and tropical cyclone (TC) forecasting.

a. Fog Cases

Enabling a forecaster to better visualize fog formation, extent, and dissipation can enhance flight safety and aid to minimize impacts from fog to busy coastal airports and for mariners at seaports. Fig. 10 shows two fog events: the top row is of a coastal fog event on September 7, 2020 with 82% illumination and the bottom row is an event off of Mexico on October 6, 2017 when there was 99% illumination. In addition to coastal fog over California on September 7, 2020, California had a series of wildfires and significant smoke. This can be seen in the Lunar Reflectance imagery as the city lights appear hazy in the interior regions near where fires were located. More details about these events are found in a recent publication by Miller et al. (2022) which also highlights the importance of reliable nighttime low cloud detection. While both cases are for periods of high illumination, the case for Mexico is outside of the initial model AOI and both show potential subjective performance over land. A quick looks shows overall greater contrast in the clear and cloud portions of the images with the best being in the DNB lunar reflectance, followed by the ML-NVI, the split window, and then traditional IR. When comparing the ML-NVI to the IR or split window, especially noticeable is how clear the extent of the fog bank appears and the texture in the ML-NVI. The ML-NVI also appears to capture to some degree, SST differences as seen in the California case. A comparison to the daily SSTs and the extent of the wild fire smoke show that the turbulent features seen in the ML-NVI imagery along the California coast must be due to the SSTs.

One of the key elements of forecasting fog dissipation is the point of the forecast location with respect to the fog boundary (Gurka 1978b,a). The ML-NVI provides good ability to detect this boundary as seen in Fig. 10 with the fog boundary for the region in the lower left corner detectable in DNB and ML-NVI but harder to detect from the IR and split window. From this one may be able to determine more precisely the extent of the fog layer and time its dissipation. Miller et al. (2022) highlighted that the bottom case was an event in which the split windows incorrectly identified fog which was then utilized as a cloud mask. It appears that this may be the case for areas along the coast and to the south west of Mexico looking at the split window vs the DNB and the ML-NVI. An additional potential use of ML-NVI can be to improve cloud masks. We also see the location of the California coastal current, compared to the open ocean in the ML-NVIs due to the significant temperature SST contrast; however, its uniform smooth texture helps to identify it as such versus low clouds for visual analysis especially when combined with the Lunar Reflectance product. There may be other beneficial uses for visualizing the SST contrast but they were not explored at this time. On land, we see city lights removed with the ML-NVI. In this case, this enhances the ML-NVIs ability to see the fog inland from the Monterrey Bay region at 122° W 37° N. Overall, the ML-NVI highlights the fog better than the IR or the split window and is helpful for detecting the fog extent when not possible with DNB due to illumination.

b. Tropical Cyclone Case

The Operational Line Scanner (OLS), the predecessor to the DNB, determined there was a 1-2 degree/ 60-120 nautical mile difference in pinpointing the center position location of tropical cyclones between using IR and nighttime visible sensors for low level circulation (Miller et al. 2006). Additionally, the turn to using nighttime visible sensors such as OLS and DNB over IR when available had significant effects on wind fields and forecast timing (Miller et al. 2006). By using the ML-NVI in the same capacity that OLS or DNB are currently used for both manual and automated processes such as in CIRA's Red-Green-Blue (RGB) multispectral enhancement or the ARCHER-II algorithm, center position fixes may become more precise than current nighttime IR fixes as illustrated previously (Wimmers and Velden 2016). Furthermore, animated ML-NVI imagery can help separate the layers of cloud rotation and may improve determination of TC intensity using Dvorak techniques and structure knowledge in comparison to LWIR animations. Examples of

this are highlighted in the image comparison for Typhoon Fengshen in Fig. 11. This cyclone case occurred outside the AOI that was used for model training and validation demonstrating an expanded use of the model.

West Pacific Typhoon Fengshen is seen in Fig 11. The samples are from November 13, 2019 at 1438 UTC (top row) and 1529 UTC (bottom row) respectively. Lunar illumination for this period was at 99%. Due to the location of the storm, successive satellite passes approximately 50 minutes apart were able to capture changes in the structure of the system. The DNB lunar reflectance in the second row shows a circulation center around 17° N and 152° E. In the corresponding ML-NVI there is a dark spot in the lower clouds in the same area, and in the IR this clearing is not present. Both rows show a long cloud feature oriented from southwest to northeast in the box bounded by 12 ° N, 14° N, 150° E and 153 ° E. The imagery contrast and structure of this feature are most prominent in the DNB lunar reflectance imagery, but is also apparent in the ML-NVI. This feature is present in the IR, but its shape is obstructed by upper level cirrus. In this case the ML-NVI provides a better assessment of storm structure than the IR for use with the Dvorak technique. Additionally, individual areas of convection centered around 18 ° N and 150 ° E are noticeable in all three imagery types but appeared to be merged in the IR while the ML-NVI permits the detection of more individual convective areas presently seen in the corresponding DNB lunar reflectance. Though this is data from the VIIRS channels, viewing two consecutive passes over the system aids to demonstrate the benefit to TC forecasting that a geostationary DNB-like product may provide even if its resolution is not as refined as the DNB sensor as instead of two images there could be from five to fifty depending on which scan is available over the storm.

c. Conclusion

In this study a ML model was developed using LWIR to create pseudo DNB lunar reflectances at full moon. Lunar reflectances are derived from measured satellite radiances enabling quantifiable metrics for nighttime synthetic imagery products. Quantitative evaluations of the ML-NVI were conducted with scores as seen in Table 3. Imagery created from the synthetic lunar reflectance provides night-time imagery to the end user that behaves similar to DNB and solar visible imagery, retains a transparent visualization of upper level clouds, and still enables the ability to see lower

level clouds unless it is a very dense and optically thick cloud as seen in Fig. 3 - Fig. 5, Fig. 7, and Fig. 10 - Fig. 11.

The implications on environmental applications are numerous. First, the ML-NVI provides a product that improves the contrast issues seen in most DNB imagery products and preserves a constant contrast regardless of the lunar cycle. This is more intuitive to a forecaster and does not require additional calculations or look up tables as may be needed for scaled DNB radiance products. This can aid in forecasting, especially for polar regions, where there is more frequent coverage with the JPSS satellites. Next, as seen above, ML-NVI can enhance the ability to detect fog and low-level tropical circulations more easily than IR alone or the low cloud split window techniques and may also aid in minimizing false alarms for cloud products. While geostationary satellites are often the preferred imagery to use for cyclone and fog forecasting, there are observing and forecasting benefits to using polar orbiter imagery for these events as well. The secondhand impacts to aviation safety are vast. These examples demonstrate the use of NVI and benefits it can bring on the JPSS systems if made available in real-time. Hillger et al. (2016) stated that the "operational applications of this nighttime imagery are the ultimate validation of its usefulness" and we have demonstrated the usefulness of the ML-NVI for low cloud identification with a focus on tropical cyclones and fog. Once the ML-NVI algorithm can be repurposed to run in near-real time, imagery can be distributed for use to a wide range of forecasters giving greater accessibility to the product.

The ML-NVI was designed as a proof of concept to demonstrate that measurements derived from IR channels common to both VIIRS and geostationary sensors, such as the ABI and Advanced Himawari Imager (AHI) can be used to create pseudo lunar reflectances like those from measured DNB radiances – in an environment where there is ample validation data to test the performance. We have demonstrated both qualitatively and quantitatively the ability to create visually consistent nighttime visible imagery from LWIR across the full lunar cycle. Ongoing research shows that the methodology and models used for JPSS VIIRS ML-NVI does translate to ABI and AHI sensors Pasillas (2024); Pasillas et al. (2023); Stanford et al. (2024). This will enable persistent nighttime visible imagery via geostationary satellites for enhanced TC monitoring and more precise visualization of fog coverage extent to aid in the timing of fog formation and dissipation over a region or site. Details on this transfer will be covered in a follow-on paper.

Qualitative Comparisons Lunar Cycle

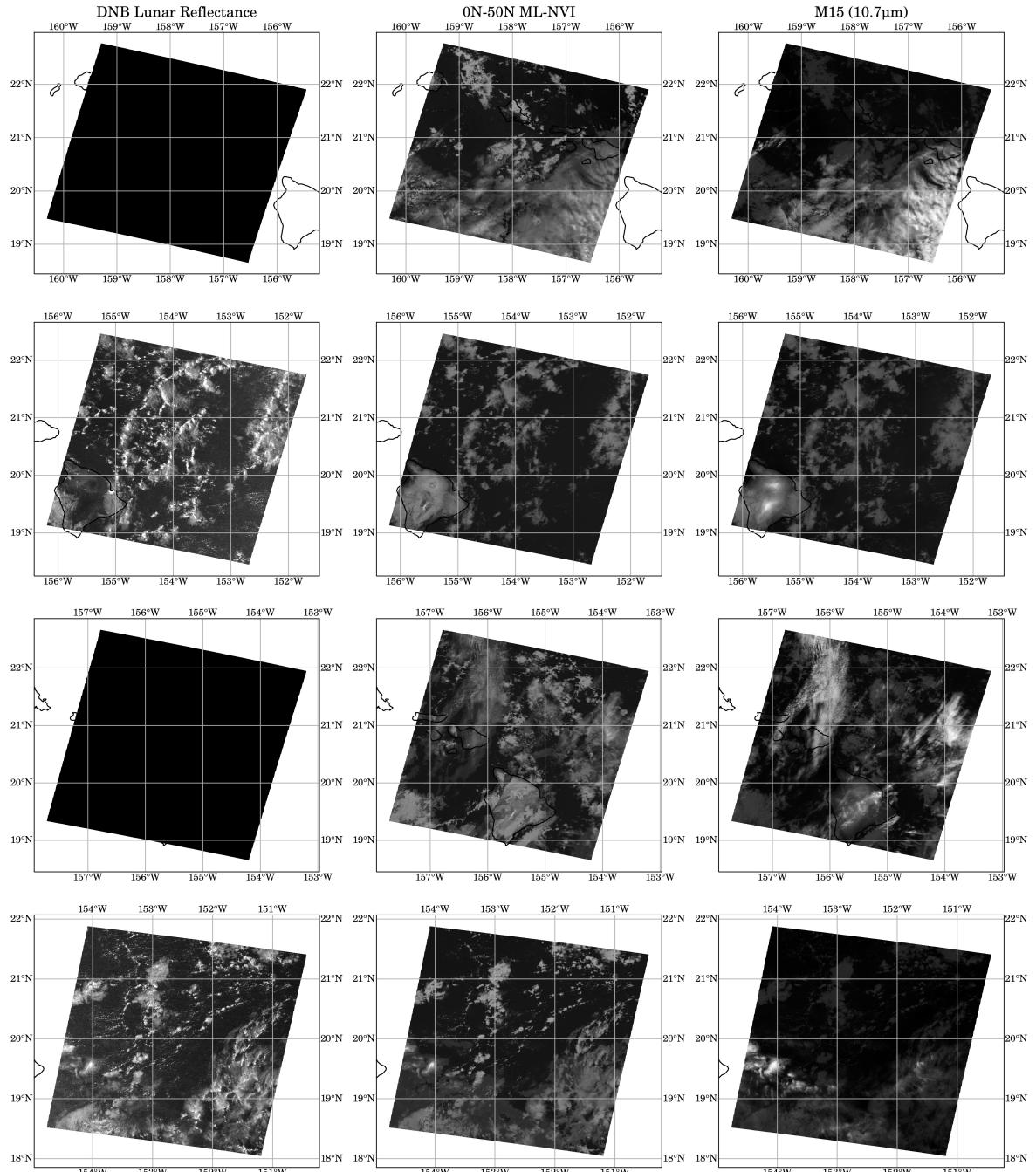


FIG. 9:

Lunar Cycle Comparisons (left to right) of DNB Lunar Reflectance, Full-latitudinal Model Lunar Reflectance, and M15 (12.01 μm) for (top to bottom) May 28, 2020 at 1210 UTC, June 2, 2020 at 1220 UTC, June 23, 2020 at 1222 UTC, and July 12, 2020 at 1128 UTC). High illumination periods are seen in rows 2 and 4 with low illumination periods in rows 1 and 3. The quality and contrast of the ML-NVI imagery seen is consistent over the full lunar cycle (center).

Qualitative Fog Comparisons

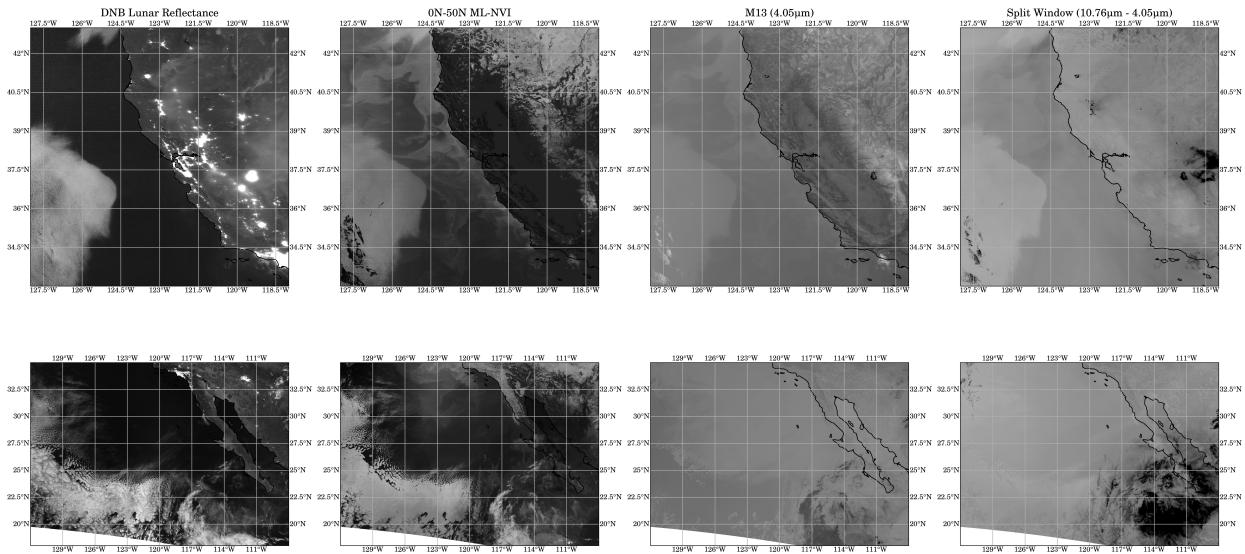


FIG. 10:

Qualitative Comparisons (Left to Right) of images from DNB radiance calculated lunar reflectance, predicted lunar reflectance for the Full latitude model, VIIRS M13 ($4.05\mu\text{m}$) and VIIRS Split Window ($10.76\mu\text{m}$ - $4.05\mu\text{m}$) sensors. Images are from the Northern Pacific open ocean on September 7, 2020 0907 UTC (top) and October 6, 2017 0931 UTC (bottom).

Qualitative Comparison for Typhoon Fengshen

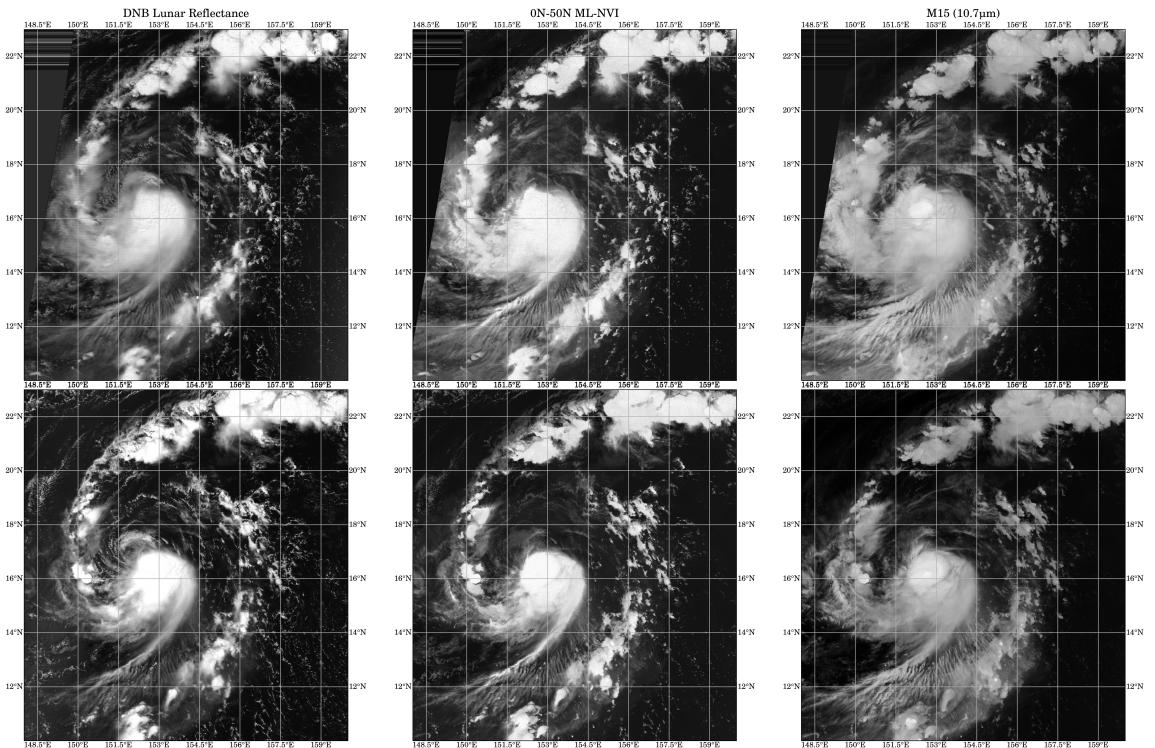


FIG. 11:

Qualitative Comparisons (Left to Right) of images from DNB radiance calculated lunar reflectance, predicted lunar reflectance for the full latitude model, and VIIRS M15 ($10.7\mu\text{m}$). Images are from two consecutive passes over Typhoon Fengshen in the west Pacific Ocean on November 13, 2019, 1438 UTC (top) and 1529 UTC (bottom).

Acknowledgments. This work was supported in part by the United States Air Force and the Office of Naval Research award N000142012069. We would like to thank Steve Reising, Kristen Rasmussen, and our anonymous reviewers for their comments.

Data availability statement. The data set for this study was downloaded from NOAA CLASS and is publicly available at (<https://www.avl.class.noaa.gov>). Code for data preparation and access to the machine learning models is available by request to the corresponding author via email or at (<https://github.com/c-pasillas>). The Miller-Turner Lunar Reflectance Model is available by request from CIRA.

References

- Ackerman, S. A., K. I. Strabala, W. P. Menzel, R. A. Frey, C. C. Moeller, and L. E. Gumley, 1998: Discriminating clear sky from clouds with modis. *Journal of Geophysical Research: Atmospheres*, **103** (D24), 32 141–32 157, <https://doi.org/10.1029/1998JD200032>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1998JD200032>, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/1998JD200032>.
- Anthis, A. I., and A. P. Cracknell, 1999: Use of Satellite Images for Fog Detection (AVHRR) and Forecast of Fog Dissipation (METEOSAT) over Lowland Thessalia, Hellas. *Int J. Remote Sensing.*, **20**:6, 1107–1124, <https://doi.org/10.1080/01431169921287>.
- Bendix, J., 2002: A Satellite-based Climatology of Fog and Low-level Stratus in Germany and Adjacent Areas. *Atmos. Res.*, **64**, 3–18, [https://doi.org/10.1016/S0169-8095\(02\)00075-3](https://doi.org/10.1016/S0169-8095(02)00075-3).
- Calvert, C., and M. Pavolonis, 2010: GOES-R Advanced Baseline Imager (ABI) Algorithm Theoretical Basis Document ATBD for Low Cloud and Fog. Algorithm Theoretical Basis Document, NOAA-NESDIS, 78 pp. <https://www.star.nesdis.noaa.gov/goesr/docs/ATBD/Imagery.pdf>.
- Chapman, R., and R. Gasparovic, 2022: *Remote Sensing Physics: An Introduction to Observing Earth from Space*. John Wiley & Sons, 496 pp.
- Chirokova, G., J. A. Knaff, and J. L. Beven, 2018: Proxy Visible Satellite Imagery. *Proc. 22nd Conference on Satellite Meteorology and Oceanography*, Austin, Texas, Amer. Meteor. Soc., 7.6, <https://ams.confex.com/ams/98Annual/webprogram/Paper334276.html>.

Chirokova, G., and Coauthors, 2023: Proxyvis—a proxy for nighttime visible imagery applicable to geostationary satellite observations. *Weather and Forecasting*, **38** (12), 2527 – 2550, <https://doi.org/10.1175/WAF-D-23-0038.1>, URL <https://journals.ametsoc.org/view/journals/wefo/38/12/WAF-D-23-0038.1.xml>.

Ellrod, G. P., 1995: Advances in the Detection and Analysis of Fog at Night Using GOES Multispectral Infrared Imagery. *Wea. Forecasting*, **10** (3), 606–619, [https://doi.org/10.1175/1520-0434\(1995\)010<0606:AITDAA>2.0.CO;2](https://doi.org/10.1175/1520-0434(1995)010<0606:AITDAA>2.0.CO;2).

Ellrod, G. P., and I. Gultepe, 2007: Inferring Low Cloud Base Heights at Night for Aviation Using Satellite Infrared and Surface Temperature Data. *Fog and Boundary Layer Clouds: Fog Visibility and Forecasting*, I. Gultepe, Ed., Birkhäuser Basel, Basel, 1193–1205, https://doi.org/10.1007/978-3-7643-8419-7_6.

Gurka, J. J., 1978a: The Role of Inward Mixing in the Dissipation of Fog and Stratus. *Mon. Wea. Rev.*, **106** (11), 1633–1635, [https://doi.org/10.1175/1520-0493\(1978\)106<1633:TROIMI>2.0.CO;2](https://doi.org/10.1175/1520-0493(1978)106<1633:TROIMI>2.0.CO;2).

Gurka, J. J., 1978b: The Use of Enhanced Visible Imagery for Predicting the Time of Fog Dissipation. Preprints, *Conference on Weather Forecasting and Analysis and Aviation Meteorology*, Washington, DC, USA, 343–346.

Harder, P., and Coauthors, 2020: NightVision: Generating Nighttime Satellite Imagery from Infra-Red Observations. *arXiv preprint arXiv:2011.07017*.

Hillger, D., and Coauthors, 2016: User Validation of VIIRS Satellite Imagery. *Remote Sensing*, **8** (1), <https://doi.org/10.3390/rs8010011>, URL <https://www.mdpi.com/2072-4292/8/1/11>.

Hillger, D. W., 2008: GOES-R Advanced Baseline Imager Color Product Development. *J. Atmos. Oceanic Technol.*, **25** (6), 853–872, <https://doi.org/10.1175/2007JTECHA911.1>.

Hoese, D., W. W. Roberts, E. Schiffer, K. Strabala, J. Feltz, R. Garcia, and J. Zeng, 2023: ssec/polar2grid: Python package version 3.0.2 (v3.0.2). <https://doi.org/10.5281/zenodo.7662308>.

Hogan, R. J., E. J. O'Connor, and A. J. Illingworth, 2009: Verification of Cloud-Fraction Forecasts. *Quart. J. of the Roy. Meteor. Soc.: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, **135** (643), 1494–1511, <https://doi.org/10.1002/qj.481>.

Inoue, T., 1989: Features of Clouds over the Tropical Pacific during Northern Hemispheric Winter Derived from Split Window Measurements. *J. of the Meteor. Soc. of Japan*, https://doi.org/10.2151/JMSJ1965.67.4_621.

Kidder, S. Q., D. W. Hillger, A. J. Mostek, , and K. J. Schrab, 2000: Two Simple GOES Imager Products for Improved Weather Analysis and Forecasting. *Natl. Weather Dig.*, **24**:4, 25–30, <http://nwafiles.nwas.org/digest/papers/2000/Vol24No4/Pg25-Kidder.pdf>.

Kidder, S. Q., and T. H. Vonder Haar, 1995: *Satellite Meteorology: an Introduction*. Gulf Professional Publishing, 480 pp.

Kieffer, H. H., and T. C. Stone, 2005: The Spectral Irradiance of the Moon. *The Astronomical J.*, **129** (6), 2887, <https://doi.org/10.1086/430185>.

Kim, H.-W., J.-M. Yeom, D. Shin, S. Choi, K.-S. Han, and J.-L. Roujean, 2017: An assessment of thin cloud detection by applying bidirectional reflectance distribution function model-based background surface reflectance using geostationary ocean color imager (goci): A case study for south korea. *Journal of Geophysical Research: Atmospheres*, **122** (15), 8153–8172, [https://doi.org/https://doi.org/10.1002/2017JD026707](https://doi.org/10.1002/2017JD026707), URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017JD026707>, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2017JD026707>.

Kim, J.-H., S. Ryu, J. Jeong, D. So, H.-J. Ban, and S. Hong, 2020: Impact of Satellite Sounding Data on Virtual Visible Imagery Generation Using Conditional Generative Adversarial Network. *IEEE J. of Selected Topics in Applied Earth Observations and Remote Sensing*, **13**, 4532–4541, <https://doi.org/10.1109/JSTARS.2020.3013598>.

Kim, K., J.-H. Kim, Y.-J. Moon, E. Park, G. Shin, T. Kim, Y. Kim, and S. Hong, 2019: Nighttime Reflectance Generation in the Visible Band of Satellites. *Remote Sensing*, **11** (18), 2087, <https://doi.org/10.3390/rs11182087>.

Kim, Y., and S. Hong, 2019: Deep Learning-generated Nighttime Reflectance and Daytime Radiance of the Midwave Infrared Band of a Geostationary Satellite. *Remote Sensing*, **11** (22), 2713, <https://doi.org/10.3390/rs11222713>.

Liang, C. K., S. Mills, B. I. Hauss, and S. D. Miller, 2014: Improved VIIRS Day/Night Band Imagery with Near-Constant Contrast. *IEEE Transactions on Geoscience and Remote Sensing*, **52** (11), 6964–6971, <https://doi.org/10.1109/TGRS.2014.2306132>.

Lin, L., and C. Cao, 2019: The Effects of VIIRS Detector-Level and Band-Averaged Relative Spectral Response Differences Between S-NPP and NOAA-20 on the Thermal Emissive Bands. *IEEE J of Selected Topics in Applied Earth Observations and Remote Sensing*, **12** (10), 4123–4130, <https://doi.org/10.1109/JSTARS.2019.2938221>.

Lindsey, D. T., L. Grasso, J. F. Dostalek, and J. Kerkmann, 2014: Use of the GOES-R split-window difference to diagnose deepening low-level water vapor. *J. Appl. Meteor. Climatol.*, **53** (8), 2005–2016, <https://doi.org/10.1175/JAMC-D-14-0010.1>.

Line, B., D. Hillger, and T. Kopp, 2018: Joint Polar Satellite System (JPSS) VIIRS Imagery Products Algorithm Theoretical Basis Document (ATBD) Revision E. Algorithm Theoretical Basis Document, NOAA NESDIS Center for Satellite Applications and Research, 55 pp.

Liou, K.-N., 2002: *An Introduction to Atmospheric Radiation*, Vol. 84. Elsevier, 608 pp.

Loew, A., and Coauthors, 2017: Validation Practices for Satellite-based Earth Observation Data Across Communities. *Reviews of Geophysics*, **55** (3), 779–817, <https://doi.org/10.1002/2017RG000562>.

Miller, S. D., J. D. Hawkins, K. Richardson, T. F. Lee, and F. J. Turk, 2006: Enhanced Tropical Cyclone Monitoring with MODIS and OLS. *Proc. 27th Hurr. Trop. Meteor. Conf. Amer. Meteor. Soc.*, Monterey, CA, Amer. Meteor. Soc., 15A.6, <https://ams.confex.com/ams/27Hurricanes/webprogram/Paper108619.html>.

Miller, S. D., and R. E. Turner, 2009: A Dynamic Lunar Spectral Irradiance Data Set for NPOESS/VIIRS Day/Night Band Nighttime Environmental Applications. *IEEE Transactions on Geoscience and Remote Sensing*, **47** (7), 2316–2329, <https://doi.org/10.1109/TGRS.2009.2012696>.

- Miller, S. D., and Coauthors, 2013: Illuminating the Capabilities of the Suomi National Polar-orbiting Partnership (NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band. *Remote Sensing*, **5** (12), 6717–6766, <https://doi.org/10.3390/rs5126717>.
- Miller, S. D., and Coauthors, 2022: A Physical Basis for the Overstatement of Low Clouds at Night by Conventional Satellite Infrared-Based Imaging Radiometer Bi-Spectral Techniques. *Earth and Space Science*, **9** (2), e2021EA002137, <https://doi.org/10.1029/2021EA002137>.
- Min, M., and Coauthors, 2017: An Investigation of the Implications of Lunar Illumination Spectral Changes for Day/Night Band-based Cloud Property Retrieval Due to Lunar Phase Transition. *JGR: Atmospheres*, **122** (17), 9233–9244, <https://doi.org/10.1002/2017JD027117>.
- Mohandoss, T., A. Kulkarni, D. Northrup, E. Mwebaze, and H. Aleomohammad, 2020: Generating Synthetic Multispectral Satellite Imagery From Sentinel-2. *arXiv preprint arXiv:2012.03108*, <https://doi.org/10.48550/arXiv.2012.03108>.
- Pasillas, C., 2024: Turning night into day: the creation, validation, and application of synthetic lunar reflectance values from the day-night band and infrared sensors for use with JPSS VIIRS and GOES ABI. Ph.D. thesis, Colorado State University, <https://mountainscholar.org/items/4f016373-f867-40e3-b58b-aad51b503cb9>.
- Pasillas, C., M. M. Bell, and C. D. Kummerow, 2023: Enhancing Low Level Closed Circulation Identification Using Night-time Visible Imagery. *Proc. 103rd Annual Meeting*, Denver, Colorado, Amer. Meteor. Soc., 15A.3, <https://ams.confex.com/ams/103ANNUAL/meetingapp.cgi/Paper/419325>.
- Pedregosa, F., and Coauthors, 2011: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830.
- Pitts, K., and M. Seybold, 2010: ABI L2 Cloud and Moisture Imagery Beta, Provisional and Full Validation Readiness, Implementation and Management Plan (RIMP) V2. RIMP, NOAA-NESDIS, 31 pp. 410-R-RIMP-0323.
- Prabhakara, C., R. S. Fraser, G. Dalu, M.-L. C. Wu, R. J. Curran, and T. Styles, 1988: Thin cirrus clouds: Seasonal distribution over oceans deduced from nimbus-4 iris. *J. Appl. Meteor.*, **27**, 379–399, [https://doi.org/10.1175/1520-0450\(1988\)027<0379:TCCSDO>2.0.CO;2](https://doi.org/10.1175/1520-0450(1988)027<0379:TCCSDO>2.0.CO;2).

Raspaud, M., and Coauthors, 2019: pytroll/satpy: Version 0.16.0. <https://doi.org/10.5281/zenodo.3250583>.

Seaman, C., D. W. Hillger, T. J. Kopp, R. Williams, S. Miller, and D. Lindsey, 2015: Visible Infrared Imaging Radiometer Suite (VIIRS) Imagery Environmental Data Record (EDR) User's Guide Version 1.3. Technical Report, NOAA-NESDIS, 55 pp. <https://doi.org/10.7289/V5/TR-NESDIS-150>, NOAA Techinal Report NESDIS, 150.

Seaman, C. J., and S. D. Miller, 2015: A Dynamic Scaling Algorithm for the Optimized Digital Display of VIIRS Day/Night Band Imagery. *Int. J. of Remote Sensing*, **36** (7), 1839–1854, <https://doi.org/10.1080/01431161.2015.1029100>.

Stanford, N. K., C. Pasillas, and A. J. Wimmers, 2024: Center Fixing Tropical Depressions and Tropical Storms Using Machine Learning - Nighttime Visible Imagery. *36th Conference on Hurricanes and Tropical Meteorology*, Long Beach, California, Amer. Meteor. Soc., 16D.3, <https://ams.confex.com/ams/36Hurricanes/meetingapp.cgi/Paper/441438>.

Strabala, K. I., S. A. Ackerman, and W. P. Menzel, 1994: Cloud Properties Inferred from 8–12- μm Data. *J. Appl. Meteor. Climatol.*, **33** (2), 212–229, [https://doi.org/10.1175/1520-0450\(1994\)033<0212:CPIFD>2.0.CO;2](https://doi.org/10.1175/1520-0450(1994)033<0212:CPIFD>2.0.CO;2).

Warren, S. G., C. J. Hahn, J. London, R. M. Chervin, and R. L. Jenne, 1988: Global Distribution of Total Cloud Cover and Cloud Type Amounts Over the Ocean. Technical Report, USDOE Office of Energy Research: Carbon Dioxide Research Div.; National Center for Atmospheric Research, 305 pp. <https://doi.org/10.2172/5415329>, URL <https://www.osti.gov/biblio/5415329>.

Weiss, M., 2001: Avn-based MOS Ceiling Height and Total Sky Cover Guidance for the Contiguous United States, Alaska, Hawaii and Puerto Rico. Technical Procedure Bulletin 483, National Weather Service, 22 pp.

Wimmers, A. J., and C. S. Velden, 2016: Advancements in objective multisatellite tropical cyclone center fixing. *J. Appl. Meteor. Climatol.*, **55**, 197–212.

Zhongming, Z., L. Linong, Y. Xiaona, Z. Wangqiang, L. Wei, and Coauthors, 2012: Recommended Methods for Evaluating Cloud and Related Parameters. Technical Report WWRP- No. 2012-1, WMO, 305 pp.

Zinke, S., 2017: A Simplified High and Near-Constant Contrast Approach for the Display of VIIRS Day/Night Band Imagery. *Int. J. of Remote Sensing*, **38** (19), 5374–5387, <https://doi.org/10.1080/01431161.2017.1338838>.