

# Temporal Interpolation of Geostationary Satellite Imagery With Optical Flow

Thomas J. Vandal<sup>id</sup> and Ramakrishna R. Nemani<sup>id</sup>

**Abstract**—Applications of satellite data in areas such as weather tracking and modeling, ecosystem monitoring, wildfire detection, and land-cover change are heavily dependent on the tradeoffs to spatial, spectral, and temporal resolutions of observations. In weather tracking, high-frequency temporal observations are critical and used to improve forecasts, study severe events, and extract atmospheric motion, among others. However, while the current generation of geostationary (GEO) satellites has hemispheric coverage at 10–15-min intervals, higher temporal frequency observations are ideal for studying mesoscale severe weather events. In this work, we present a novel application of deep learning-based optical flow to temporal upsampling of GEO satellite imagery. We apply this technique to 16 bands of the GOES-R/Advanced Baseline Imager mesoscale dataset to temporally enhance full-disk hemispheric snapshots of different spatial resolutions from 10 to 1 min. Experiments show the effectiveness of task-specific optical flow and multiscale blocks for interpolating high-frequency severe weather events relative to bilinear and global optical flow baselines. Finally, we demonstrate strong performance in capturing variability during convective precipitation events.

**Index Terms**—Image processing, multispectral imaging, neural networks, remote sensing, scientific computing.

## I. INTRODUCTION

EVERY second, satellites around the earth are generating valuable data to monitor weather, land cover, infrastructure, and human activity. Satellite sensors capture reflectance/radiance intensities at designated spectral wavelengths, spatial, and temporal resolutions. Properties of the sensors, including wavelengths and resolutions, are optimized for particular applications. Most commonly, satellites are built to capture the visible wavelengths, which are essentially RGB images. Scientific specific sensors capture a larger range of wavelengths, such as micro, infrared, and thermal waves, providing information to many applications at temporal frequencies from 1 min at mesoscale to multiple years at climate scales. Satellites carrying sensors typically follow one of two orbits, geostationary (GEO) and polar. Polar-orbiting satellites

most often have daily or longer revisit times and include NASA's A-Train [1], Landsat-8 [2], and Sentinel [3]. Data provided by the Moderate Resolution Imaging Spectroradiometer (MODIS) [4] carried on Terra/Aqua and Landsat are widely used for quantifying effects of climate change, land-cover usage, and air pollution, among others, but are not well suited to monitoring high-frequency events. On the other hand, GEO satellites are well suited for subdaily events, such as tracking weather and studying diurnal cycles. While a high altitude reduces spatial resolution, current generation multispectral sensors on GEO satellites are able to provide 1–15-min observations, enabling immense opportunity for understanding atmospheric, land cover, and oceanic dynamics. High temporal frequency observations from satellites are critical for studying extreme environmental events, such as storm tracking and wildfire detection, but, due to physical constraints, are not available at global scales. In this work, we present an approach to temporally interpolate global 10-min observations to 1 min by learning from spatially isolated mesoscale observations.

Within a few years, a constellation of GEO satellites by multiple international institutions will provide global coverage of the earth's state. The latest generation of GEO satellites includes NOAA/NASA's GOES-16/17 [5], Japan's Himawari-8/9 [6], China's Fengyun-4 [7], and Korea's GEO-KOMPSAT-2A with future plans in development. Full-disk coverage from such satellites has revisit times of 10–15 min allowing applications to real-time detection and observation of wildfires [8], hurricane tracking, air quality, flooding, precipitation estimation, flood risk, and others [9]. Furthermore, given improved spectral and spatial resolutions in current-generation sensors, GEO satellites open future opportunities to incorporate and learn from less frequent observations from polar orbiters.

While 10–15-min revisit times are temporally sufficient for many applications, higher frequency snapshots can aid a variety of tasks. For instance, understanding rapidly evolving convective events is a high priority for improving atmospheric models, which are notoriously poor at simulating heavy precipitation, as highlighted in NASA's Earth Science Decadal Survey [10]. However, data for analyzing such events are often not available at the desired frequency. Similarly, comparing multiple satellite observations is dependent on their corresponding timestamps. This leads to an interpolation task between observations in a multispectral spatiotemporal sequence, similar to that of video interpolation.

Temporal interpolation requires a weighted combination of two images. The simplest approach to interpolation is

Manuscript received January 4, 2021; revised June 11, 2021; accepted July 24, 2021. This work was supported in part by the NASA Earth eXchange (NEX) and in part by the NASA Research Opportunities in Space and Earth Science (ROSES) under Grant 19-ESROGSS-19-0123. (Corresponding author: Thomas J. Vandal.)

Thomas J. Vandal is with the NASA Ames Research Center, Moffett Field, CA 94035 USA, and also with the Bay Area Environmental Research Institute, Moffett Field, CA 94035 USA (e-mail: vandat@baeri.org).

Ramakrishna R. Nemani is with the NASA Ames Research Center, Moffett Field, CA 94035 USA.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TNNLS.2021.3101742>.

Digital Object Identifier 10.1109/TNNLS.2021.3101742

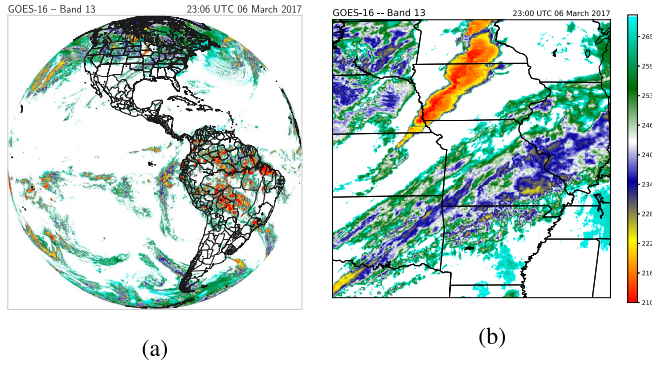


Fig. 1. Cloud-top temperature (K) from (a) “clean” IR long-wave window band 13 full-disk and (b) mesoscale coverages.

to directly weigh each pixel in the image by the relative difference in time. Modern techniques to this problem use deep learning and optical flow to warp images based on movement captured by tracking apparent motion [11]–[13]. In this work, we present a novel application of deep learning-based optical flow to the problem of temporal interpolation between GEO satellite imagery. We compare properties of global and spectral specific optical flow interpolation using the SuperSloMo method (SSM) [13] with baselines. Our experiments show that spectral-specific SSM is capable of interpolating high-frequency atmospheric events. Furthermore, visual analysis suggests that the learned flow resembles atmospheric motion and dynamic visibility maps.

The remainder of this article is outlined as follows. Section II discusses related work including resolution enhancement in the earth sciences and the current state of video frame interpolation. Section III introduces the GOES-R dataset. Section IV details methodology. Experiments on a large-scale dataset and a severe storm case study are presented in Section V. Finally, Section VI concludes with challenges and future work.

## II. GOES-R SATELLITE DATASET

GEO satellites are synchronized in orbit with the earth’s spin to hover over a single location. Given this location, the sensor, measuring radiation as often as possible, can frequently capture data over a continuous and large region. This feature makes GEO satellites ideal for capturing environmental dynamics. The GOES-R series satellites, namely, GOES-16/17 (east and west sides of the Americas), operated by NASA and NOAA provide scientists with unprecedented temporal frequency enabling real-time environmental monitoring using the Advanced Baseline Imager (ABI) [14]. GOES-16/17 senses 16 bands of data that are viewed in Fig. 2 and listed in Table I with central wavelength, spatial resolution, and band name. Three data products are derived from each GOES-16/17: 1) full disk covering the western hemisphere every 10 min (15 min prior to April 2, 2019) [see Fig. 1(a)]; 2) Continental U.S. every 5 min; and 3) mesoscale user-directed 1000 km  $\times$  1000 km subregion every at an optimal 30 s [see Fig. 1(b)]. ABI’s 16 spectral bands include two visible (1 and 2), four near-infrared (3–6), and ten infrared (7–16) bands enabling a suite of applications.

TABLE I  
GOES-R SERIES BANDS

Band	Central Wavelength ( $\mu\text{m}$ )	Spatial Resolution (km)	Name
1	0.47	1	Blue
2	0.64	0.5	Red
3	0.86	1	Veggie
4	1.37	1	Cirrus
5	1.6	1	Snow/Ice
6	2.24	2	Cloud Particle Size
7	3.9	2	Shortwave Window
8	6.2	2	Upper-level Water Vapor
9	6.9	2	Mid-level Water Vapor
10	7.3	2	Low-Level Water Vapor
11	8.4	2	Cloud-Top Phase
12	9.6	2	Ozone
13	10.3	2	“Clean” IR Longwave
14	11.2	2	IR Longwave
15	12.3	2	“Dirty” IR Longwave
16	13.3	2	CO <sub>2</sub> Longwave IR

These GEO satellites are particularly useful in tracking weather, monitoring high-intensity events, estimating rainfall rates, fire detection, and many others in near real time. The mesoscale mode gives forecasters the ability to “point” the satellite at a user-specific subregion for near-constant monitoring of severe events. For example, GOES-16 and GOES-17 actively provide emergency response units tools for decision-making during wildfires in the Western United States. These high-frequency data provide valuable information on environmental dynamics and retrospective analysis, such as studying convective events [15]. Furthermore, mesoscale data can be used to inform techniques to produce higher temporal resolution CONUS and full-disk coverage. In this work, we develop a model to improve the temporal resolutions of CONUS and full disk by learning an optical flow model to interpolate between consecutive frames. With this, we are able to generate 1-min full-disk artificially enhanced data.

## III. RELATED WORK

In this section, we begin by reviewing previous work in the areas of data fusion and resolution enhancement as applied generally to remote sensing satellite imagery and some recent successes of deep learning in the area. Second, we provide a brief review of optical flow and video frame interpolation techniques.

### A. Resolution Enhancement of Satellite Data

Earth science datasets are complex and often require extensive preprocessing and domain knowledge to effectively render themselves useful for large-scale applications or monitoring. Such datasets may contain frequent missing values due to sensor limitations, low-quality pixel intensities, incomplete global coverage, and contamination with atmospheric processes related to cloud and aerosols. Furthermore, spatial and temporal resolution enhancement is often applied to improve analysis precision. Techniques to handle these challenges have been developed and are widely applied across the remote sensing community.

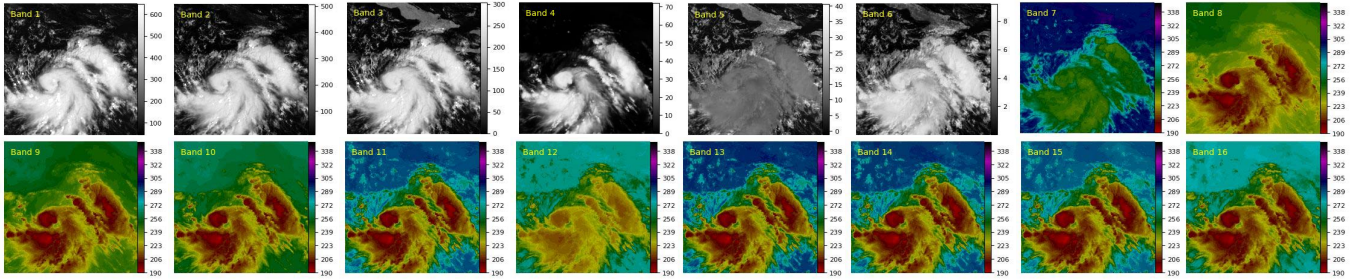


Fig. 2. 16 bands observed by GOES-16 mesoscale, including two visible, four near-infrared, and ten thermal infrared. Observation from mode RadM1 on October 5, 2020, 15:05 UTC.

Many statistical and machine learning methodologies for improving spatial resolution have been explored and are an active area of research. Data fusion is one area where two or more datasets are *fused* to generate an enhanced product, often with both higher spatial and temporal resolutions [16]. The Spatial and Temporal Adaptive Reflectance Fusion (STARFM) algorithm, for example, uses Landsat and MODIS to produce a daily 30-m reflectance product by using a spectralwise weighting model [17]. Using CubeSats, Houborg and McCabe [18] present an approach to further leverage Landsat and MODIS for further spatiotemporal improvements. Similarly, nearest neighbor analog multiscale patch-decomposition data-driven models are used as state-of-the-art interpolation techniques for developing global sea surface temperature (SST) datasets [19]. In recent years, super-resolution techniques have presented state-of-the-art results for spatial enhancement of satellite images [20]–[22].

Approaches for temporal resolution enhancement of individual satellite observations have not been as well studied. Liebmann and Smith [23] presented the first linearly interpolated datasets filling in missing and erroneous long-wave radiation many days apart to improve global coverage. Similarly, Kandasamy *et al.* [24] presented a comparison of multiple methods for interpolating between MODIS observations to generate a synthetic leaf area index dataset. A number of statistical techniques, including long-term climatology measures and time-series decomposition, were applied to smooth observation and fill gaps. Reference [25] presented an approach using linear interpolation on subdaily GEO imagery to match timestamps between multiple satellites. However, given more frequent observations by the recent generation of GEO observations, more complex methods beyond linear interpolation may be more applicable and accurate in the temporal domain.

Our work proposes to apply deep learning-based optical flow methodologies to optimize the interpolation problem. In recent years, a number of applications in processing and learning from satellite data have shown state-of-the-art results using deep learning. For example, Benedetti *et al.* [26] showed that recurrent and convolutional neural networks effectively assimilate multiple satellite images. Lanaras *et al.* [22] presented a global deep learning super-resolution approach for Sentinel-2 with a 50% improvement beyond traditional techniques. In terms of classification, DeepSat showed that normalized deep belief networks that are tuned were able to outperform traditional techniques for image classifica-

tions [27]. Convolutional neural networks have been shown to effectively classify land use in remotely sensed images, from urban areas [28] to crop types [29].

While many studies have explored resolution enhancement spatially, and temporally, the authors are not aware of any prior work exploring temporal interpolation at the minute-to-minute scale. Prior approaches on longer time scales have applied linear interpolation and nearest-neighbor techniques. We explore the applicability of a more complex optical flow approach to temporal interpolation at very high resolutions and use linear interpolation as our baseline, as applied in prior work. Atmospheric motion vectors (AMVs) are a related line of work focused on tracking the movement of clouds and corresponding heights to initialize numerical weather prediction and data assimilation models [30], [31]. Traditional techniques to AMVs use spatial similarities across time to approximate lateral movement and often produce low yields and coarsened resolution, which do not apply to temporal interpolation. Recently, the optical flow has been shown to be effective at this task [32], which could motivate future work.

### B. Optical Flow and Video Frame Interpolation

Temporally interpolating between frames of images can be computed with a weighted average of two frames, warped or not, to a defined intermediate time,  $t$ . Linear interpolation is the simplest approach and can be written as  $I(t) = I_0 * t + I_1 * (1 - t)$ , given two frames  $I_0$  and  $I_1$ . However, this technique fails to account for any motion, physical phenomena, and corresponding occlusion that occur between two frames giving a poor performance. Temporal interpolation techniques applied to video have shown high skill at generating slow-motion footage by generating intermediate frames in spatially and temporally coherent sequences [11]–[13], [33]. These approaches are designed to learn the dynamics by inferring displacement of spatial structure between consecutive images. Optical flow is widely used for this task, which estimates spatial displacement by comparing movement between two images.

In recent years, deep learning architectures have shown promising results for both optical flow and video interpolation. Supervised learning of optical flow is often constrained by the availability of training data as motion is rarely quantified in real images. Datasets, such as Flying Chairs [34] and MPI Sintel [35], have been generated synthetically and are used for methods development but may not be realistic to real-world scenes. High-performing architectures for supervised optical



flow learning include FlowNet [36], [37], PWC-Net [38], and RAFT [39] that take advantage of encoder-decoder, volume-based correlations, pyramid, and recurrent layers. However, since optical flow labels are rarely available, as is the case in our GEO temporal interpolation, unsupervised learning techniques are often applied to real-world naturally generated datasets. Unflow [40] presented an approach to training FlowNet using a bidirectional occlusion-aware unsupervised reconstruction loss. Recent studies have improved unsupervised learning further, such as Geonet [41] with depth perception and [42] by predicted occlusion directly. In practice, we have found that training fully unsupervised optical flow networks on satellite imagery to be a nontrivial exercise providing poor performance in our unreported experiments, likely due to physical complexity in the dataset.

Many video interpolation techniques focus on *single frame* interpolation, meaning that a single frame is estimated by the model directly between two consecutive frames at  $t = 0.5$  [11], [33], [43]. However, when interpolating satellite imagery, time-dependent and multiframe estimation is preferred for more flexibility. Jiang *et al.* [13] presented SSM that combines both optical flow and occlusion models for time-dependent estimation between consecutive frames. The time-dependent nature of this approach produces spatially and temporally coherent predictions of any time between 0 and 1. In their experiments, Jiang *et al.* [13] show that 240-frames/s video clips can be estimated from 30-frames/s inputs. Further details of this work will be presented in Section IV where we apply their architecture with an extension to multiscale optical flows.

High-frequency satellite imagery can take advantage of these techniques to extract the dynamics of different physical processes. Our application requires a time-dependent methodology with no labels available for supervised optical flow learning. Furthermore, our imagery has 16 spectral bands in the visible, near-infrared, and thermal-infrared spectra to capture physical phenomena and with varying spatial resolution. Between frames, these processes, such as convection, break optical flow's consistency assumption making unsupervised learning a challenge. SSM provides a fundamental technique to temporal interpolation that is well suited to our application. We generalize this approach with task-specific models for each spectral band in our dataset. In the remainder of this work, we study how SSM can be effectively applied to this problem by experimenting with global- and task-specific models.

#### IV. METHODOLOGY

Temporal upsampling of GEO satellite data is a similar problem as intermediate video frame interpolation with domain-specific characteristics. In video interpolation, the goal is to estimate an intermediate frame given two or more consecutive images. A single set of optical flows is sufficient for interpolating between RGB images as objects captured in the visible spectrum are reasonably consistent across frames. However, as discussed above, satellite imagery often consists of tens or even hundreds of spectral channels with varying spatial resolutions. Furthermore, each channel captures different physical properties with heterogeneous motion, including

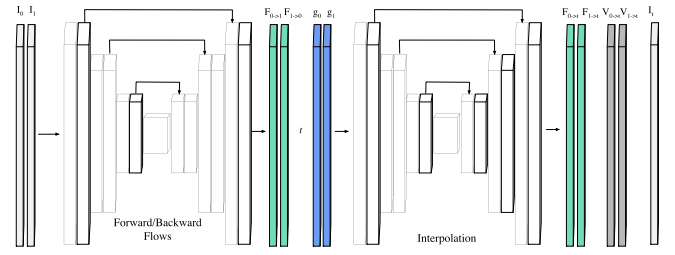


Fig. 3. SuperSloMo architecture with flow and interpolation networks.

severe events, such as convection leading to heavy precipitation and tornadoes. The goals of the proposed methodologies include interpolating to a user-defined point in time, capturing varying spatial dynamics, and computational efficiency at scale. In this section, we describe the SSM framework [13] for temporal upsampling with the optical flow with our domain-specific adaptations with global- versus task-specific networks. The global model performs interpolation on all channels simultaneously, while task specific can be learned for each channel independently.

##### A. Intermediate Frame Interpolation

SSM intermediate frame interpolation considers the case of frame estimation at a user-defined point in continuous time [13]. To ensure smooth transitions and structural similarity between frames, SSM is designed to predict optical flows between two input images as a function of time. The approach, which can be seen in Fig. 3, consists of two deep neural networks. The first estimates forward and backward flows between two input images. The second network, depending on time, updates the forward and backward flows and generates visibility maps to handle occlusion. These features of SSM are well suited to GEO data by enabling arbitrary temporal upsampling and synchronization of multiple datasets.

Let  $I_0, I_1, I_t \in \mathcal{R}^{H \times W \times C}$ , where  $t \in (0, 1)$ ,  $H$  is the image height,  $W$  is the image width, and  $C$  is the number of spectral bands. *Task-specific* optical flow is defined when  $C = 1$ . The goal is then to construct an intermediate frame  $I_t$  with a linear combination of warped  $I_0$  and  $I_1$  as defined by

$$\hat{I}_t = \alpha \cdot g(I_0, F_{0 \rightarrow t}) + (1 - \alpha) \cdot g(I_1, F_{t \rightarrow 1}) \quad (1)$$

where  $F_{0 \rightarrow t}$  and  $F_{t \rightarrow 1}$  are the optical flows from  $I_0$  to  $I_t$  and  $I_t$  to  $I_1$ , respectively.  $g$  is defined as the *backward warping* function, implemented with bilinear interpolation, and  $\alpha$  represents a scalar weight coefficient to enforce temporal consistency and allow for occlusion reasoning. In the case of high temporal resolution satellite imagery, the interpolation is virtually estimating the state of atmospheric variables (clouds, water vapor, and so on) over a static land surface. If a given pixel in  $I_0$  captures land surface, but the same pixel in  $I_1$  sees a cloud, the occlusion principle is used to estimate at what time  $t$  the cloud covers the pixel. Furthermore, atmospheric dynamics cause physical characteristics to change over time. One example is convection such that warm/cold air vertically and rapidly mixes in the atmosphere causing severe weather events. In the context of interpolating, dynamics between  $I_0$  and  $I_1$  cause

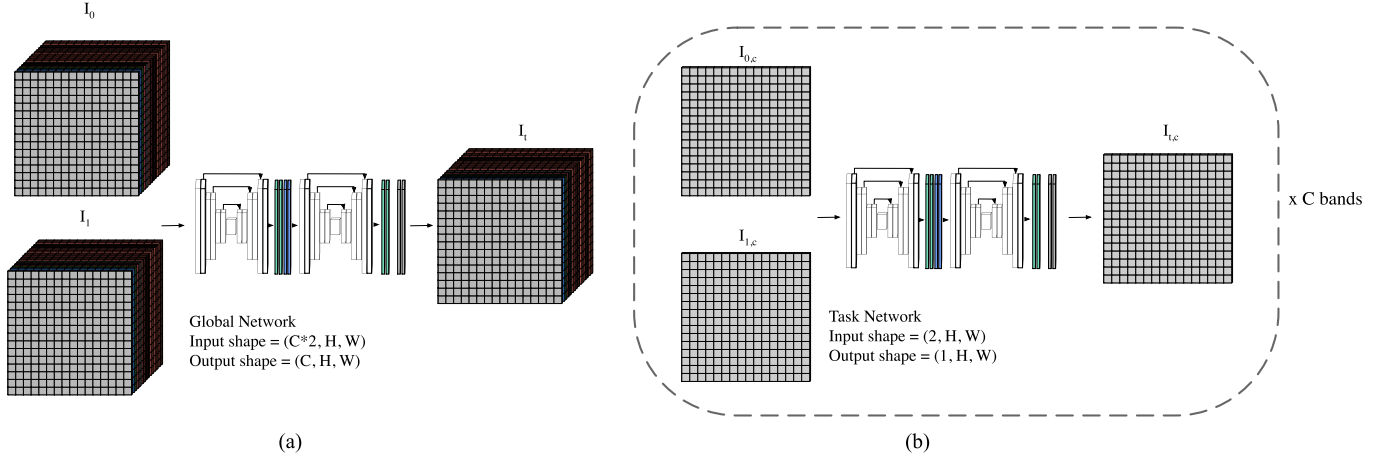


Fig. 4. Global-versus task-specific networks for interpolation on multispectral cubes with SSM. (a) SSM-G. (b) SSM-T.

cloud temperature to rapidly decrease, leading to a drastic change brightness intensity and breaking consistency assumptions of optical flow. Visibility maps,  $V_{0 \rightarrow t}, V_{1 \rightarrow t} \in (0, 1)^{H \times W}$ , weight brightness importance to account for both occlusion and intensity changes. Equation 1 is then be redefined as

$$\hat{I}_t = \frac{1}{Z} \cdot ((1 - t) \cdot V_{0 \rightarrow t} \cdot g(I_0, F_{0 \rightarrow t}) + t \cdot V_{1 \rightarrow t} \cdot g(I_1, F_{1 \rightarrow t})) \quad (2)$$

where  $Z = (1 - t) \cdot V_{0 \rightarrow t} + t \cdot V_{1 \rightarrow t}$  is a normalization factor. Forward and backward optical flows,  $(F_{0 \rightarrow t}, F_{1 \rightarrow t})$ , at time  $t$  are estimated by a sequence of two flow networks,  $G_{\text{flow}}$  and  $G_{\text{Interp}}$ , as presented in Fig. 3. The first network,  $G_{\text{flow}}(I_0, I_1)$ , infers backward and forward optical flows,  $(F_{0 \rightarrow 1}, F_{1 \rightarrow 0})$ , between two input images. After generating approximate intermediate flows,  $(\hat{F}_{0 \rightarrow t}, \hat{F}_{1 \rightarrow t})$ , intermediate images are generated. The interpolation network,  $G_{\text{Interp}}$ , predicts visibility maps  $(V_{0 \rightarrow t}, V_{1 \rightarrow t})$  and final flows  $(F_{0 \rightarrow t}, F_{1 \rightarrow t})$  as a function of a concatenation of input images, intermediate flows, and intermediate warped images.

### B. Task-Specific Interpolation

Multispectral satellite imagery has multiple spectral channels, each observing varying phenomena, such as clouds moving faster at higher levels of the atmosphere. The movement of objects within images of varying spatial resolutions can have dramatic effects on the performance of optical flow networks. In traditional optical flow, interpolation features are tracked using a single model for single or three-channel images in the visible range. The optical flow assumption of brightness consistency is relatively well satisfied on the pixel level in high frame-rate sequences. In satellite images, different movements appear in each channel of the data with underlying physical processes affecting brightness intensity. Rather than modeling all the channels in a single SSM model, we propose to model each channel independently. We denote *SSM-G* as the global SSM model where all channels are modeled simultaneously, as presented in Fig. 4. *SSM-T<sub>c</sub>* denotes task-specific networks that are trained for individual channel  $c$ . Formally, SSM-T

learns a set of SSM models  $[I_t^c]$  for  $c \in C$ . While requirements for graphics processing unit (GPU) computation multiply with SSM-T, we will show that improved performance of task-specific models improves results substantially.

### C. Network Architecture

Deep neural networks with encoding and decoding are well suited to model both local and global spatial structures. Architectures of this type include FlowNet [36] and U-Net [13], which have been shown to perform well in the task of optical flow. We follow this approach using a U-Net architecture for each of the flow and interpolation networks. The U-Net architecture applied has four downsampling layers followed by four upsampling layers with skip connections between each corresponding layer. A convolution layer maps the input to 64 channels with a kernel size of 7. The following downsampling layers are of size 128, 256, 512, and 512 with kernel sizes 5, 5, 3, and 3. Each downsampling layer performs average pooling and two convolutions with rectified linear unit (ReLU) activations. Upsampling layers of sizes 256, 128, 64, and 32 all with kernel sizes of 3 are then applied. Each layer performs bilinear interpolation followed by two convolutions with ReLU activations. Finally, 32 channels in the last hidden layer are mapped to the number of output channels using a convolution operation of kernel size 3. Flow and interpolation networks use the same architecture with different input and output dimensions as discussed above.

Tracking both small and large displacements continues to be a challenge, even with encoder-decoder network architectures. Other approaches have shown that using a stack of networks performing small and large displacements performs well [37]. In this work, we explore the applicability of multiscale hidden layers to track local and global features. We follow a similar approach applied in [44] where hidden layers are defined to have multiple convolution operations with different sized kernels followed by a concatenation layer. In our networks, kernels of sizes 3, 5, and 7 conserve high-frequency spatial details while abstracting global motion for improved optical flows and visibility maps.

#### D. Training Loss

As all variables in the architecture are differentiable, the model can be learned in an end-to-end manner. Given two inputs frames  $I_0$  and  $I_1$  with  $N$  intermediate frames  $\{I_{t_i}\}_{i=1}^N$  and corresponding predictions  $\{\hat{I}_{t_i}\}_{i=1}^N$ , a loss function can be defined as a weighted combination of reconstruction, warping, and smoothness losses such that

$$l = \lambda_r l_r + \lambda_w l_w + \lambda_s l_s. \quad (3)$$

We note that Jiang *et al.* [13] include a fourth term for the perception of image classes that are not available for this satellite dataset. Similarly, we employ  $L_1$  loss functions for each loss term unless noted otherwise.

The *reconstruction loss* is defined as the distance between observed and predicted intermediate frames

$$l_r = \frac{1}{N} \sum_{i=1}^N \|\hat{I}_{t_i} - I_{t_i}\|. \quad (4)$$

A *warping loss* is used to optimize estimated optical flows between input and intermediate frames for channel  $c$

$$\begin{aligned} l_w = & \|I_0 - g(I_1, F_{0 \rightarrow 1})\| + \|I_1 - g(I_0, F_{1 \rightarrow 0})\| \\ & + \frac{1}{N} \sum_{i=1}^N \|I_{t_i} - g(I_0, F_{0 \rightarrow t_i})\| \\ & + \frac{1}{N} \sum_{i=1}^N \|I_{t_i} - g(I_1, F_{1 \rightarrow t_i})\|. \end{aligned} \quad (5)$$

A *smoothness loss* is applied to forward and backward flows from  $I_0$  to  $I_1$  to satisfy the smoothness assumption of optical flows in the first network such that

$$l_s = \|\nabla F_{0 \rightarrow 1}\|_1 + \|\nabla F_{1 \rightarrow 0}\|_1. \quad (6)$$

In practice, this training setup requires optimization over multiple hyperparameters, including  $\lambda_r$ ,  $\lambda_s$ ,  $\lambda_w$ , and a learning rate.

#### V. EXPERIMENTS

We demonstrate the effectiveness of a set of SSM models on a large-scale dataset using a high-performance computing system with a cluster of GPUs. The goal of our experiments is to show that optical flow is highly applicable for temporal interpolation of satellite imagery and compare to the baseline of linear interpolation, as traditionally applied. Section V-A outline the training process, compare methodologies, and study the effectiveness of a severe convective precipitation event. Code for this work can be found in the Supplementary Material.<sup>1</sup>

##### A. Training

Data for training and testing were taken from the GOES-16 Mesoscale 1-min imagery. These images are of identical spatial and spectral resolution as North America and full-disk imagery, so the learned models are directly applicable to

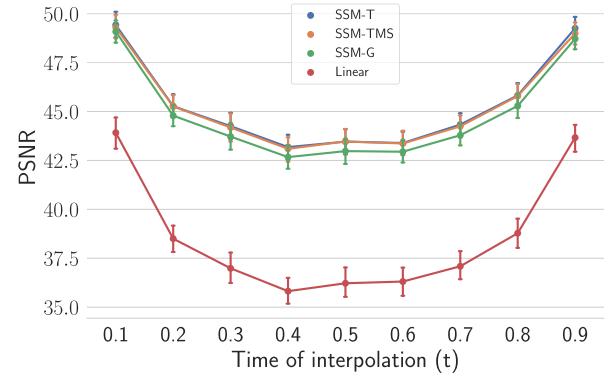


Fig. 5. Interpolation error as a function of time.

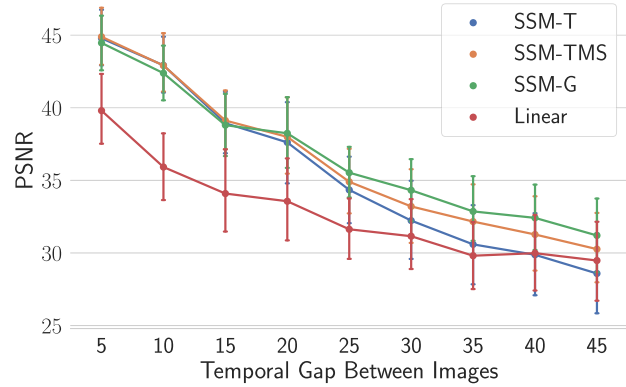


Fig. 6. Interpolation error as a function of temporal gap between observations.

these datasets. Training data were selected using all samples for every five days of the year 2018 and testing data on a randomly selected set of examples from 2019. Samples were generated as  $264 \times 264$  subimages and randomly cropped to  $256 \times 256$  during training. Standardized normalization was applied independently to each channel to ensure similar pixel intensity distributions across bands. Temporally, samples are selected from a sequence of 15 time steps such that inputs ( $I_0, I_1$ ) are 10-min apart with a random label  $I_t$  in-between. Furthermore, during training, images are randomly flipped and rotated to improve generality in the U-Net architecture. A random training/validation split of 20% was used to monitor learning. We select cloud top temperature tracked by band 13 ( $10.3 \mu\text{m}$ ) in ablation and demonstration experiments as used in studies of convection and AMVs. Experiments for this study leveraged NASA's Pleiades Supercomputer and the NASA Earth eXchange (NEX) to process large-scale GOES-16 data and train individual networks for each of the 16 channels.

Adam optimization is used to minimize (3) with default parameters  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ , and  $\text{eps} = 1\text{e-}8$  in PyTorch. We found that learning is sensitive to hyperparameters  $\lambda_s$  and  $\lambda_w$  and are optimized using probabilistic grid search and constrained Bayesian optimization [45]. Constrained Bayesian optimization applies efficient randomized Monte Carlo simulations over  $\lambda_s$  and  $\lambda_w$  holding  $\lambda_r = 1$ . We perform this process using the open-source Ax library [46]

<sup>1</sup><https://github.com/tjvandal/geostationary-superslomo>

TABLE II  
MODEL COMPARISON RESULTS FROM 200 RANDOMLY SELECTED SAMPLES IN 2019. **BOLD HIGHLIGHTS**  
THE TOP PERFORMING MODEL, AND \* HIGHLIGHTS THE SECOND BEST

Band	$\mu\text{m}$	PSNR $\uparrow$				RMSE $\downarrow$				SSIM $\uparrow$			
		Linear	SSM-G	SSM-T	SSM-TMS	Linear	SSM-G	SSM-T	SSM-TMS	Linear	SSM-G	SSM-T	SSM-TMS
1	0.47	37.795	36.828	<b>38.282</b>	37.818*	0.178*	0.198	<b>0.160</b>	0.185	0.719	0.682	<b>0.734</b>	0.722*
2	0.64	37.408	37.006	<b>37.748</b>	37.583*	0.185	0.186	<b>0.169</b>	0.177*	0.637	0.616	<b>0.649</b>	0.644*
3	0.86	<b>41.808</b>	40.544	41.350*	41.135	0.100*	0.112	<b>0.099</b>	0.108	<b>0.760</b>	0.712	0.731	0.737*
4	1.37	60.519	60.838	<b>62.598</b>	61.925*	0.012	0.011	<b>0.008</b>	0.009*	0.969	0.974	<b>0.983</b>	0.982*
5	1.6	<b>56.097</b>	55.129	56.044*	55.703	<b>0.018</b>	0.019	<b>0.018</b>	<b>0.018</b>	0.932	0.928	<b>0.937</b>	0.935*
6	2.24	55.076	58.316	58.693*	<b>58.758</b>	0.373	0.255	<b>0.242</b>	<b>0.242</b>	0.747	0.884	0.893*	<b>0.895</b>
7	3.9	40.721	46.084	<b>46.591</b>	46.496*	1.825	0.972	<b>0.917</b>	0.932*	0.766	0.899	<b>0.907</b>	0.905*
8	6.2	50.669	57.656	<b>58.432</b>	58.135*	0.613	0.374	<b>0.226</b>	0.358*	0.747	0.907	<b>0.913</b>	0.912*
9	6.9	47.476	55.287	56.014*	<b>56.015</b>	0.904	0.336	0.306*	<b>0.305</b>	0.756	0.924	<b>0.929</b>	<b>0.929</b>
10	7.3	44.601	52.535	<b>53.226</b>	53.120*	1.222	0.582	<b>0.418</b>	0.550*	0.748	0.919	<b>0.924</b>	<b>0.924</b>
11	8.4	38.530	44.753	45.184*	<b>45.243</b>	2.335	1.071	<b>1.020</b>	1.013*	0.770	0.922	<b>0.929</b>	<b>0.929</b>
12	9.6	43.560	49.568	50.023*	<b>50.030</b>	1.314	0.626	<b>0.594</b>	<b>0.594</b>	0.762	0.913	<b>0.921</b>	0.920*
13	10.3	38.667	44.925	<b>45.439</b>	45.343*	2.286	1.177	<b>0.991</b>	1.130*	0.782	0.925	<b>0.933</b>	0.932*
14	11.2	38.167	44.594	44.996*	<b>45.090</b>	2.392	1.080	1.036*	<b>1.024</b>	0.770	0.924	0.931*	<b>0.932</b>
15	12.3	38.163	44.699	<b>45.284</b>	45.252*	2.387	1.185	<b>0.991</b>	1.118*	0.762	0.921	0.929*	<b>0.930</b>
16	13.3	40.578	47.313	47.731*	<b>47.778</b>	1.819	0.786	0.751*	<b>0.745</b>	0.721	0.892	0.898*	<b>0.899</b>

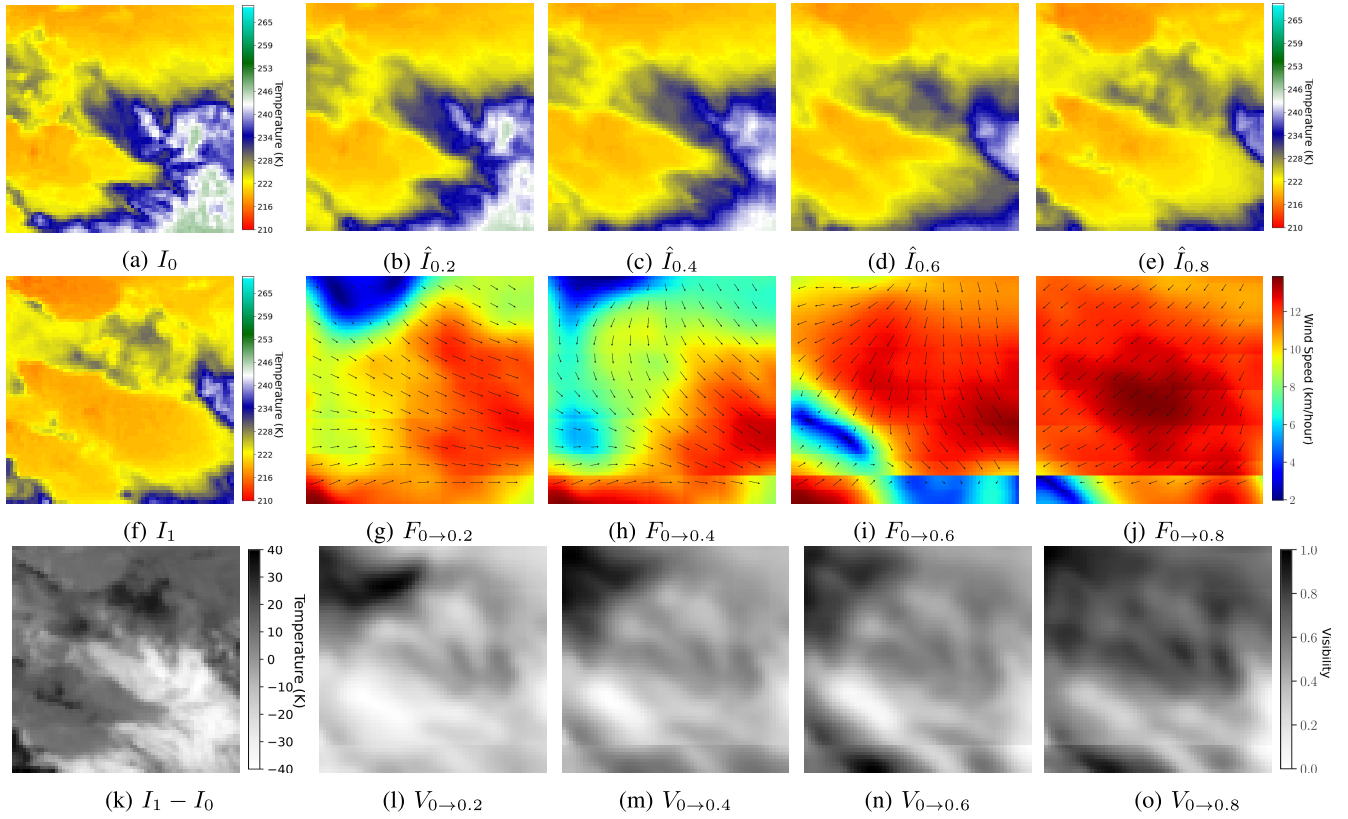


Fig. 7. Severe convective event on May 23, 2019 from 2:00 to 2:30 UTC taken from GOES-16 mesoscale band 13, clean long-wave infrared ( $10.3 \mu\text{m}$ ). (a) and (f) are in the input images, and their difference in 7(k). (b)–(e) SSM-T interpolated predictions, flow intensity, and direction in (g)–(j) and visibility maps in (l)–(o).

for 20 trials on band 1 with SuperSlomo and find  $\lambda_s = 0.23$  and  $\lambda_w = 0.65$  minimized reconstruction loss on the validation set. These hyperparameters are applied to all following bands and experiments. Training the suite of models was executed on a multinode GPU cluster of Nvidia V100's with one model per band, each taking one day to train. Multi-GPU training was used for serial hyperparameter optimization and global models.

### B. Model Comparison

This section compares variations of SuperSlomo with a linear interpolation baseline for interpolation of GEO images. Linear interpolation between frames is performed by taking a linear combination of two input images weighted by time,  $\hat{I}_t = (1-t) * I_0 + t * I_1$ . A set of three SuperSlomo models are explored, including global (SSM-G), task specific (SSM-T), and task specific with multiscale layers (SSM-TMS). SSM-T



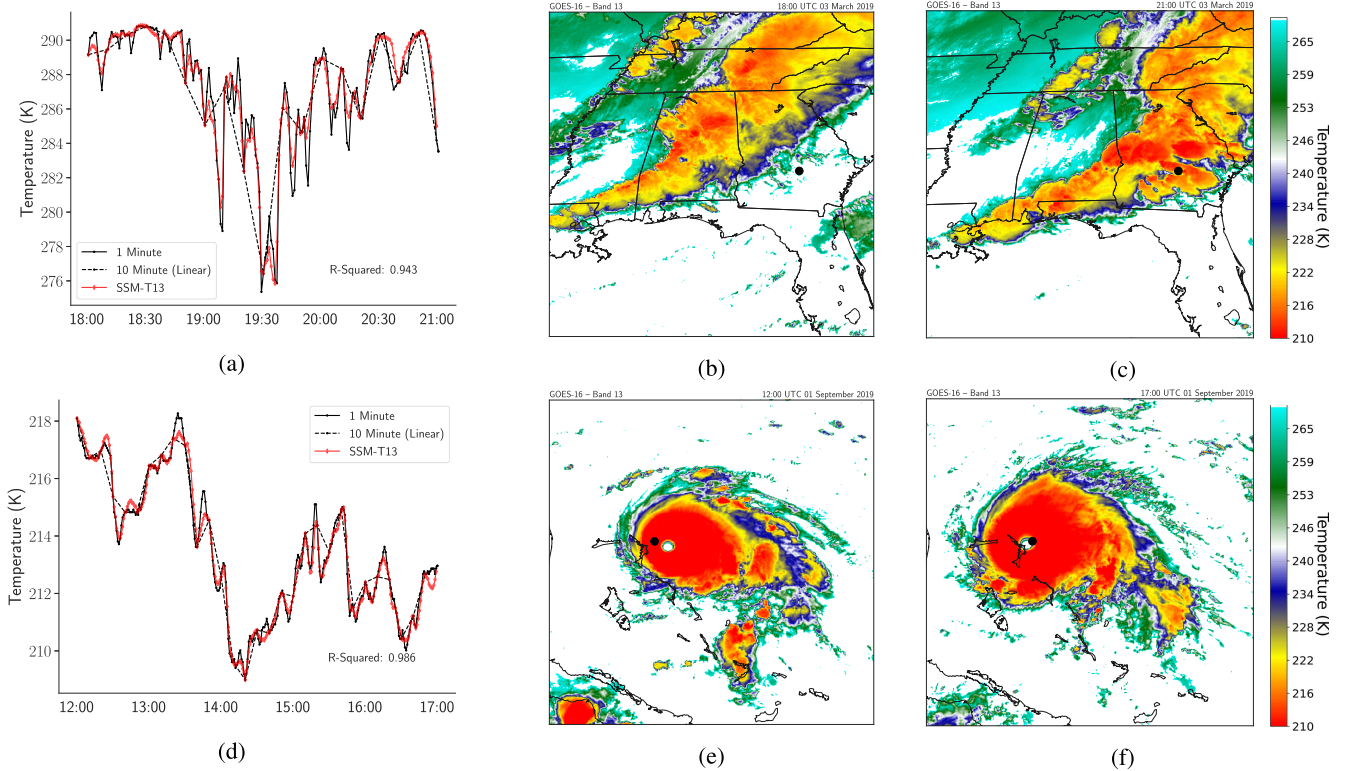


Fig. 8. Time-series interpolation of long-wave clean IR band 13 ( $10.3 \mu\text{m}$ ) during two severe weather events. Left to right: observed and interpolated time series followed by starting and ending observations. (a)–(c) Observation of  $32.0^\circ$  latitude and  $-83.0^\circ$  longitude on March 3, 2019, from 18:00 to 21:00 UTC, depicted on images with a red circle. (d)–(f) Observation of  $26.7^\circ$  latitude and  $-76.8^\circ$  longitude on September 1, 2019, from 12:00 to 17:00 UTC.

models are trained for each band separately. SSM-G is trained using training data from all bands and, hence, a substantially larger training set. The root mean square error (RMSE), the peak-to-signal-noise ratio (PSNR), and the self-similarity measure (SSIM) are used to evaluate performance.

We first study the inherent properties of SSM on band 13, including time dependence and sensitivity to larger displacements. Interpolation between two frames is expected to have smooth transitions from one frame to another. Generally, interpolation will have the largest error where the distance to frames is maximum (i.e., directly between the input frames). In Fig. 5, we compare PSNR as a function of  $t \in [0, 1]$  between models and see this effect. The gap between linear and SSM models is pronounced. Between SSM models, SSM-T and SSM-TMS have similar performance. SSM-G that is a more generalized model does not perform quite as well as SSM-T and SSM-TMS, suggesting task-specific models across bands perform better. Fig. 6 shows PSNR at  $t = 0.5$  while increasing the gap between  $I_0$  to  $I_1$  from 5 to 45 min. A 45-min gap contains 9x more displacement than a 5-min gap making the optical flow problem more difficult. Over the first 15 min, SSM models perform similarly and better than linear. As the gap widens, SSM-TMS and SSM-G begin performing better than SSM-T. This suggests that SSM-TMS multiscale layers may be capturing more motion. SSM-G’s more diverse dataset includes 500-m data, which has larger displacements than the 2-km band 13.

In Table II, we present the results for each of the 16 bands of GOES-16 in 200 randomly sampled 10-min intervals from

2019. Metrics are computed for each sample at  $t = 0.5$ , where the error is the largest and averages over all examples. Errors are in terms of brightness temperatures measured by the sensor. As a whole, our results find that task-specific SSM models, SSM-T and SSM-TMS, outperform linear interpolation and a single global interpolation network, SSM-G. Interpolation of the visible (1 and 2) bands with optical flow provides modest improvements in all metrics. Performance improvements are not found for NIR bands 3 and 5, veggie, and snow/ice, respectively, where each method produces high PSNRs and low RMSEs. On the other hand, thermal bands find large improvements from SSM optical flow, increasing SSIM from approximately 0.77 to 0.93, on average. We find that task-specific models outperform the global model throughout even with the reduced training data size. Errors do vary across bands, which is largely associated with the radiance distribution of that particular band. For instance, band 4 ( $1.37 \mu\text{m}$ ) is typically sparse such that predicting low intensities is an easy task.

### C. Severe Weather Event

This section studies an example of two convective precipitation events visualized in Figs. 7 and 8. In the context of severe weather, convection is vertical motion in the atmosphere that occurs when warm air on the surfaces forces cold air in the atmosphere down often causing supercells and heavy precipitation. For the first time, Apke *et al.* [47] studied this process using GOES-14 1-min imagery for a set of supercells. The authors found that atmospheric motion can help



define signatures of supercell events to better inform weather forecasting models. Here, we show that cloud top brightness from the long-wave clean IR band can be interpolated from 10 to 1 min during two convective events.

The 1-min mesoscale (M1) data from May 23, 2019, from 2:00 to 3:00 UTC at  $-95^\circ$  longitude and  $37^\circ$  latitude are used for the analysis. In this region, a convective storm is occurring and moving east. The data are downsampled to 10 min interpolated back to a 1-min time series. Fig. 7(a) shows the region of interest with predictions ( $I_t$ ), optical flows ( $F_{0 \rightarrow t}$ ), and visibility maps ( $V_{0 \rightarrow t}$ ) between times 0 and  $t$ . The optical flows show the storm moving east and slightly rotating with a maximum displacement around the storm edges. Visibility pixels correspond to edges of clouds, which allows (1) to be a nonlinear combination relative to time. The flows and visibility maps in Fig. 7(h)–(j) and (m)–(o) show an increasingly apparent horizontal strip of contrasting flows caused by two artifacts of the approach. First, the approach uses both flows and occlusion maps to take a nonlinear combination of beginning and end images and can weigh some areas more than others. Second, U-Net depends on constant input size, so images are interpolated in blocks with stride 20, which can cause artifacts. As our approach is not optimizing flows directly, but rather the interpolation, these artifacts are less pronounced in the predictions.

Fig. 8 presents time series of two severe events with the corresponding observations. The first row depicts a tornado outbreak in the southeastern United States on March 3, 2019, where we see variability in cloud top brightness. A dashed line shows the 10-min time series and is equivalent to linear interpolation. SSM-T overlayed the observation and well captures the variability of a drastic 15 K temperature dropout interpolation approach that generates a highly correlated time series with an R-squared of 0.955. The second row shows a time series through the eye of Hurricane Dorian on September 1, 2019. Similarly, we find that SSM-T is highly correlated with an R-squared of 0.986. These results suggest that optical flow may be a promising approach for interpolating GEO imagery for applications to severe events.

## VI. CONCLUSION

This work proposes that temporal interpolation with the optical flow is capable of modeling high-frequency events between GEO images with high accuracy by learning from mesoscale rapid-scan observations. Experiments showed that learning independent weights of SSM for each band improve performance beyond one global SSM and the linear interpolation baseline. Multiscale blocks in SSM-TMS perform well for larger displacements and are comparable to SSM-T overall. Interpolation well captured temporal variability of cloud top brightness during multiple severe convective events. This interpolation has direct applications to improve precipitation estimation and weather variability. Furthermore, deep learning-based optical flow routines are able to better harness graphical processing units with a single feedforward pass [36] relative to traditional and more expensive routines, such as the polynomial expansion method applied in [32]. Improvements

and accuracy from deep learning approaches to optical flow present a promising direction for future work.

While further analysis is necessary, our results suggest that dynamics of atmospheric motion are learned by the network using displacement flows and visibility maps, which would have direct implications for weather forecasting. Second, internal dynamics captured may provide knowledge on how to predict future states as applied for video-frame prediction. In future work, we will explore the accuracy of optical flow to estimating atmospheric motion relative to large-scale observations and model interpretability to better understand which physical dynamics are captured.

## ACKNOWLEDGMENT

The authors would like to thank the GOES-R data providers for access to GOES-16 mesoscale and full-disk observations. NASA's Advanced Supercomputing Facility at the Ames Research Center was used for accelerated computing with Nvidia V100 GPUs.

## REFERENCES

- [1] G. L. Stephens *et al.*, "The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation," *Bull. Amer. Meteorol. Soc.*, vol. 83, no. 12, pp. 1771–1790, 2002.
- [2] D. P. Roy *et al.*, "Landsat-8: Science and product vision for terrestrial global change research," *Remote Sens. Environ.*, vol. 145, pp. 154–172, Apr. 2014.
- [3] R. Torres *et al.*, "GMES Sentinel-1 mission," *Remote Sens. Environ.*, vol. 120, pp. 9–24, May 2012.
- [4] T. S. Pagano and R. M. Durham, "Moderate resolution imaging spectroradiometer (MODIS)," *Proc. SPIE Sensor Syst. Early Earth Observing Syst. Platforms*, vol. 1939, pp. 2–18, Aug. 1993.
- [5] *Goes-R Advanced Baseline Imager (ABI) Algorithm Theoretical Basis Document for Cloud and Moisture Imagery Product (CMIP)*. Accessed: Jun. 24, 2019. [Online]. Available: <https://www.star.nesdis.noaa.gov/goesr/docs/ATBD/Imagery.pdf>
- [6] K. Bessho, K. Date, M. Hayashi, A. Ikeda, T. Imai, and H. Inoue, "An introduction to Himawari-8/9—Japan's new-generation geostationary meteorological satellites," *J. Meteorol. Soc. Japan*, vol. 94, no. 2, pp. 151–183, 2016.
- [7] J. Yang, Z. Zhang, C. Wei, F. Lu, and Q. Guo, "Introducing the new generation of Chinese geostationary weather satellites, Fengyun-4," *Bull. Amer. Meteorol. Soc.*, vol. 98, no. 8, pp. 1637–1658, Aug. 2017.
- [8] G. Xu and X. Zhong, "Real-time wildfire detection and tracking in Australia using geostationary satellite: Himawari-8," *Remote Sens. Lett.*, vol. 8, no. 11, pp. 1052–1061, Nov. 2017.
- [9] T. J. Schmit, S. S. Lindstrom, J. J. Gerth, and M. M. Gunshor, "Applications of the 16 spectral bands on the advanced baseline imager (ABI)," *J. Oper. Meteorol.*, vol. 6, no. 4, pp. 33–46, Jun. 2018.
- [10] S. S. Board, *Thriving Our Changing Planet: A Decadal Strategy for Earth Observation From Space*. New York, NY, USA: Academic, 2019.
- [11] Y. L. Liu, Y. T. Liao, Y. Y. Lin, and Y. Y. Chuang, "Deep video frame interpolation using cyclic frame generation," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, 2019, pp. 8794–8802.
- [12] W. Bao, W.-S. Lai, C. Ma, X. Zhang, Z. Gao, and M.-H. Yang, "Depth-aware video frame interpolation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 3703–3712.
- [13] H. Jiang, D. Sun, V. Jampani, M.-H. Yang, E. Learned-Miller, and J. Kautz, "Super SloMo: High quality estimation of multiple intermediate frames for video interpolation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Dec. 2018, pp. 9000–9008.
- [14] T. J. Schmit, P. Griffith, M. M. Gunshor, J. M. Daniels, S. J. Goodman, and W. J. Lebar, "A closer look at the ABI on the GOES-R series," *Bull. Amer. Meteorol. Soc.*, vol. 98, no. 4, pp. 681–698, 2017.
- [15] T. Fiolleau and R. Roca, "An algorithm for the detection and tracking of tropical mesoscale convective systems using infrared images from geostationary satellite," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4302–4315, Jul. 2013.

- [16] D. L. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proc. IEEE*, vol. 85, no. 1, pp. 6–23, Jan. 1997.
- [17] X. Zhu, J. Chen, F. Gao, X. Chen, and J. G. Masek, "An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions," *Remote Sens. Environ.*, vol. 114, no. 11, pp. 2610–2623, 2010.
- [18] R. Houborg and M. F. McCabe, "A cubesat enabled spatio-temporal enhancement method (CESTEM) utilizing planet, landsat and MODIS data," *Remote Sens. Environ.*, vol. 209, pp. 211–226, May 2018.
- [19] R. Fablet, P. H. Viet, and R. Lguensat, "Data-driven models for the spatio-temporal interpolation of satellite-derived SST fields," *IEEE Trans. Comput. Imag.*, vol. 3, no. 4, pp. 647–657, Dec. 2017.
- [20] D. Yang, Z. Li, Y. Xia, and Z. Chen, "Remote sensing image super-resolution: Challenges and approaches," in *Proc. IEEE Int. Conf. Digit. Signal Process. (DSP)*, Jul. 2015, pp. 196–200.
- [21] F. Li, L. Xin, Y. Guo, D. Gao, X. Kong, and X. Jia, "Super-resolution for GaoFen-4 remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 1, pp. 28–32, Jan. 2018.
- [22] C. Lanarasa, J. Bioucas-Dias, S. Galliani, E. Baltsavias, and K. Schindler, "Super-resolution of sentinel-2 images: Learning a globally applicable deep neural network," *ISPRS J. Photogram. Remote Sens.*, vol. 146, pp. 305–319, Dec. 2018.
- [23] B. Liebmann and C. A. Smith, "Description of a complete (interpolated) outgoing longwave radiation dataset," *Bull. Amer. Meteorol. Soc.*, vol. 77, no. 6, pp. 1275–1277, 1996.
- [24] S. Kandasamy, F. Baret, A. Verger, P. Neveux, and M. Weiss, "A comparison of methods for smoothing and gap filling time series of remote sensing observations—application to MODIS LAI products," *Biogeosciences*, vol. 10, no. 6, pp. 4055–4071, Jun. 2013.
- [25] D. R. Doelling *et al.*, "Geostationary enhanced temporal interpolation for CERES flux products," *J. Atmos. Ocean. Technol.*, vol. 30, no. 6, pp. 1072–1090, 2013.
- [26] P. Benedetti, D. Ienco, R. Gaetano, K. Ose, R. G. Pensa, and S. Dupuy, "M<sup>3</sup>fusion: A deep learning architecture for multiscale multimodal multitemporal satellite data fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 12, pp. 4939–4949, Oct. 2018.
- [27] S. Basu, S. Ganguly, S. Mukhopadhyay, R. DiBiano, M. Karki, and R. Nemani, "DeepSat: A learning framework for satellite imagery," in *Proc. 23rd SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst.*, Nov. 2015, p. 37.
- [28] M. Castelluccio, G. Poggi, C. Sansone, and L. Verdoliva, "Land use classification in remote sensing images by convolutional neural networks," 2015, *arXiv:1508.00092*. [Online]. Available: <http://arxiv.org/abs/1508.00092>
- [29] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 778–782, May 2017.
- [30] K. M. Bedka and J. R. Mecikalski, "Application of satellite-derived atmospheric motion vectors for estimating mesoscale flows," *J. Appl. Meteorol.*, vol. 44, no. 11, pp. 1761–1772, Nov. 2005.
- [31] C. Velden, W. E. Lewis, W. Bresky, D. Stettner, J. Daniels, and S. Wanzong, "Assimilation of high-resolution satellite-derived atmospheric motion vectors: Impact on HWRF forecasts of tropical cyclone track and intensity," *Monthly Weather Rev.*, vol. 145, no. 3, pp. 1107–1125, Mar. 2017.
- [32] Q. Wu, H.-Q. Wang, Y.-J. Lin, Y.-Z. Zhuang, and Y. Zhang, "Deriving AMVs from Geostationary satellite images using optical flow algorithm based on polynomial expansion," *J. Atmos. Oceanic. Technol.*, vol. 33, no. 8, pp. 1727–1747, Aug. 2016.
- [33] Z. Liu, R. A. Yeh, X. Tang, Y. Liu, and A. Agarwala, "Video frame synthesis using deep voxel flow," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4463–4471.
- [34] M. Aubry, D. Maturana, A. A. Efros, B. C. Russell, and J. Sivic, "Seeing 3D chairs: Exemplar part-based 2D-3D alignment using a large dataset of CAD models," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 3762–3769.
- [35] D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black, "A naturalistic open source movie for optical flow evaluation," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2012, pp. 611–625.
- [36] A. Dosovitskiy *et al.*, "FlowNet: Learning optical flow with convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 2758–2766.
- [37] E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox, "FlowNet 2.0: Evolution of optical flow estimation with deep networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, p. 6.
- [38] D. Sun, X. Yang, M.-Y. Liu, and J. Kautz, "PWC-net: CNNs for optical flow using pyramid, warping, and cost volume," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8934–8943.
- [39] Z. Teed and J. Deng, "RAFT: Recurrent all-pairs field transforms for optical flow," in *Proc. Eur. Conf. Comput. Vis.* Glasgow, U.K.: Springer, 2020, pp. 402–419.
- [40] S. Meister, J. Hur, and S. Roth, "Unflow: Unsupervised learning of optical flow with a bidirectional census loss," in *Proc. 32nd AAAI Conf. Artif. Intell.*, 2018, pp. 7251–7259.
- [41] Z. Yin and J. Shi, "Geonet: Unsupervised learning of dense depth, optical flow and camera pose," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Dec. 2018, pp. 1983–1992.
- [42] Y. Wang, Y. Yang, Z. Yang, L. Zhao, P. Wang, and W. Xu, "Occlusion aware unsupervised learning of optical flow," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4884–4893.
- [43] S. Niklaus, L. Mai, and F. Liu, "Video frame interpolation via adaptive separable convolution," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 261–270.
- [44] L. Ding, K. Zhao, X. Zhang, X. Wang, and J. Zhang, "A light-weight U-Net architecture multi-scale convolutional network for pediatric hand bone segmentation in X-ray image," *IEEE Access*, vol. 7, pp. 68436–68445, 2019.
- [45] B. Letham, B. Karrer, G. Ottoni, and E. Bakshy, "Constrained Bayesian optimization with noisy experiments," *Bayesian Anal.*, vol. 14, no. 2, pp. 495–519, Jun. 2019.
- [46] Ax. Accessed: Jun. 20, 2019. [Online]. Available: <https://github.com/facebook/Ax>
- [47] J. M. Apke, J. R. Mecikalski, and C. P. Jewett, "Analysis of mesoscale atmospheric flows above mature deep convection using super rapid scan geostationary satellite data," *J. Appl. Meteorol. Climatol.*, vol. 55, no. 9, pp. 1859–1887, Sep. 2016.



**Thomas J. Vandal** received the B.S. degree in mathematics from the University of Maryland at College Park, College Park, MD, USA, in 2012, and the Ph.D. degree in interdisciplinary engineering from Northeastern University, Boston, MA, USA, in 2018.

He is currently a Research Scientist with the NASA Ames Research Center, Moffett Field, CA, USA, and the Bay Area Environmental Research Institute, Moffett Field. His research interests include machine learning and earth science applications.



**Ramakrishna R. Nemani** received the B.S. degree from Andhra Pradesh Agricultural University, Bapatla, India, in 1979, the M.S. degree from Punjab Agricultural University, Ludhiana, India, in 1982, and the Ph.D. degree from the University of Montana, Missoula, MT, USA, in 1987.

He is currently a Senior Earth Scientist with the Ecological Forecasting Laboratory, NASA Ames Research Center, Moffett Field, CA, USA. His research interests include ecological forecasting and collaborative computing in the earth sciences.