Ensemble-Guided Tropical Cyclone Track Forecasting for Optimal Satellite Remote Sensing

Vinay Ravindra[ID], *Member, IEEE*, Sreeja Nag[ID], *Member, IEEE*, and Alan Li[ID]

Abstract—Within the realm of satellite remote sensing, optimal data acquisition to study natural phenomena under time, resource, and cost constraints is a well-known problem. Furthermore, since the sensors themselves are at remote locations with sparse ground connectivity, the optimal method must use a computationally light forecasting algorithm, which assimilates information from the observations at possibly irregular intervals, in near real time. In this article, we propose and demonstrate the ensemble -guided cyclone track forecasting (EGCTF) method for application in remote tropical cyclone tracking. The algorithm uses ensemble data produced by numerical weather prediction models to guide the forecasting process while assimilating measured cyclone center positions. The algorithm was tested and analyzed with the Global Ensemble Forecasting System (GEFS) data and the National Hurricane Center data for the 2018 year hurricanes within the Atlantic basin. Compared with a baseline method that uses the GEFS-issued mean ensemble track (AEMN) for forecasting and no data assimilation, the proposed algorithm exhibited positive forecast skill for more than 290 test cases over forecast periods spanning 6–48 h. The skill is seen to improve with lengthening forecast periods, with five test cases showing greater than 75% skill for a forecast period of 6 h to 247 test cases for the forecast period of 48 h.

Index Terms—Forecasting, optimization, remote sensing, satellite constellations, tropical cyclones (TCs).

I. INTRODUCTION

SATELLITE-based remote sensing has enabled new discoveries and increased our knowledge of natural phenomena governing our planet. Large amounts of data containing spectral information of our planet are downloaded and analyzed daily for widespread scientific use. In particular, predictive models are able to assimilate these data sets to increase their accuracy and prediction capabilities regarding highly impactful phenomena, such as tropical cyclones (TCs), floods, and wildfires.

While traditional Earth observation (EO) satellites have been massive, expensive, and monolithic in nature, over the past decade, there has been significant growth of small-sized

satellite systems demonstrating meaningful scientific applications. For example, the Cyclone Global Navigation Satellite System (CYGNSS) constellation of eight microsatellites launched in December 2016 weighs only 24.7 kg each and operates at a power budget of 38 W while measuring GPS signals reflected from the Earth's surface [1]. When CYGNSS data were assimilated into a research version of the hurricane weather research and forecasting (HWRF) model, a numerical weather prediction (NWP) model, improved track, intensity, and structure forecasts for hurricanes Harvey and Irma (2017) were made [2]. RainCube demonstrated an active miniaturized Ka-band precipitation profiling radar [3] and TEMPEST-D demonstrated a millimeter-wave radiometer measuring five frequencies [4], both in a 6U Cubesat form factor (launched May 2018). Thyphoon Trami (2018) was co-observed by both RainCube and TEMPEST-D to yield horizontal and vertical profiles of the typhoon [3]. Since small satellite platforms have demonstrated useful scientific data gathering, the next step is in commissioning of large-scale constellation missions with small satellites. Efforts in this field have already been started with the CYGNSS [1], Temporal Experiment for Storms and Tropical Systems (TEMPEST) [4], and Time-Resolved Observations of Precipitation structure and storm Intensity with a Constellation of Smallsats (TROPICS) missions. TROPICS is aimed at providing microwave measurements of 3-D temperature and humidity over the tropical region, as well as cloud ice and precipitation horizontal structure with a constellation of six 3U Cubesats in three orbital planes [5].

While satellites can take advantage of larger geographical scope and access to remote areas on Earth, they have to operate under constraints of a tight power budget, limited communication opportunities, bandwidth, and limited computational resources. This is especially restrictive for the case of small-sized EO satellites. Low Earth orbit (LEO) satellites are limited by access time over regions of interest and large gaps between accesses over a given region much more than satellites at higher orbits. Satellite-to-ground communications are limited by the number and positions of available ground stations. For satellites operating with active radars, the instrument duty cycle is limited by power and thermal considerations. For an efficient allocation of resources, the satellite should ideally focus its time and resources at making "interesting or useful" observations, as opposed to the case where the satellite only makes nominal observations that are filtered out later on the ground to yield interesting/useful data. This is especially true for the aforementioned radars or high-resolution imagers that

can maneuver to a large field of regard but capture within a small field of view [6]. Furthermore, since access to ground control centers is limited, the decisions would ideally need to be made autonomously onboard the satellite in scenarios when the natural phenomena under observation (such as TCs) are evolving at a rapid rate. For example, Nag *et al.* [7] described a case study of a satellite constellation monitoring urban floods with autonomous optimal scheduling and delay-tolerant networking (DTN) protocol for intersatellite communications.

Autonomous satellite tasking of taking observations of high priority events based on acquired and processed onboard imagery was demonstrated by the Earth observation-1 (EO-1) spacecraft (mass of 588 kg, launch November 2000) through the autonomous sciencecraft experiment (ASE). Active volcano eruptions were detected, prioritized, and observed [8]. The spacecraft detected flooded areas by finding the increase in the water covered pixels from acquired images and autonomously retasked the spacecraft to take additional scenes (including those regions upstream or downstream along the river to monitor the extent of the flooding) [9]. The data collected by the satellite were filtered, and only the useful data were downlinked to ground stations demonstrating huge savings in downlink resources.

A fundamental component of the onboard decision-making process would be for the satellite to learn and predict the natural phenomena, adapting its strategy based on its successive observations and that of other satellites within the constellation. Given that the computational resources on the satellite are limited, the algorithm must be computationally lightweight. In this article, we propose and investigate a new algorithm guided by ensemble outputs of NWPs to predict the track of TCs, by considering the prior ensemble prediction tracks of an NWP and the history of observations. By predicting the future positions of the TC, the satellites can dedicate its resources and optimize its limited access time over the area of TC activity to procure scientific measurements of higher value.

This article is arranged as follows. Section II describes related research and highlights the challenges in their application in autonomous onboard implementation. Sections III and IV describe the proposed algorithm and the results from the application of the algorithm for the case of the 2018 year TCs in the Atlantic basin. Section V concludes with a discussion.

II. RELATED WORK AND MOTIVATION

Roy and Kovordányi [10] provided an extensive review of different TC forecasting techniques. The accounting of the recent-past behavior of the current cyclone and/or the behavior of previously encountered similar cyclones is common to all forecasting techniques. The underlying models are based on either physical governing equations of fluid motion (dynamical models solved using numerical techniques also referred to as NWP models), statistical models using historically recorded data of TCs and the current TC (solved using regression, artificial neural networks (ANNs), Markov chains, and so on), or a hybrid of the above (statistical–dynamical models that use predictors based on the execution of the dynamical models).

Dynamical model techniques have been shown to yield increasingly accurate forecasts as evidenced in the National Hurricane Center (NHC) report for the 2018 hurricane season [11]. However, dynamical models have high computational requirements and require execution on super-computers to produce predictions within a reasonable time frame. On the other hand, statistical models have relatively low computational requirements but suffer from relatively lower forecast accuracy [12]. There have been efforts in employing machine learning on past data of TCs and/or satellite images of the current TC to produce track forecasts at near real time. Kovordányi and Roy [13] showed that the TC track direction can be determined based on cloud patterns associated with the TC using neural networks applied to satellite images from the National Oceanic and Atmospheric Administration (NOAA)—Advanced Very High Resolution Radiometer (AVHRR) instrument aboard NOAA's polar-orbiting satellite. Rüttgers *et al.* [14] used generative adversarial networks (GANs) applied to geosynchronous (GEO) satellite images and velocity fields, to produce a 6-h advance tracks for TCs. Moradi Kordmahalleh *et al.* [15] used sparse recurrent neural networks (RNNs) with dynamic time warping for predictions of hurricanes at 6- and 12-h time frames from the current time of observation, using a database of similar past hurricane trajectories.

Unfortunately, application of such forecasting techniques onboard a satellite is not straightforward. Implementation of an onboard dynamical model that directly assimilates the data observed would be computationally prohibitive. Furthermore, dynamical models require a wide range of inputs (such as wind, precipitation, and temperature), which may not be available to the satellite from its observations. For techniques that use neural networks, predicted TC positions are produced at fixed intervals from the current observation time and also expect for the input data to be available at predefined temporal intervals. This is typically available for GEO imagery since GEO satellites have a fixed view of a large region of Earth, and thus, images within this area can be potentially taken at regularly cadenced intervals. For example, the Advanced Baseline Imager (ABI) onboard the Geostationary Operational Environmental Satellite-R (GOES-R) series of GEO satellites possess a mesoscale operation mode in which coverage is provided over a 1000×1000 km^2 geographical area with a temporal resolution of 30 s and a spatial resolution of 0.5–2 km [16]. LEO satellites, however, can make observations only at aperiodic intervals. Therefore, both the input observation data and the output predictions are processed at irregular time intervals.

With the aforementioned considerations, the requirements of an onboard-satellite forecast algorithm can be outlined as follows.
1) Must impose low computational load for the satellite processor.
2) Able to accept information from observations at irregular intervals and also be able to produce forecasts for arbitrary times.
3) Limited dependence upon specific natural phenomena attributes (such as temperature and precipitation) since

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

RAVINDRA *et al.*: ENSEMBLE-GUIDED TC TRACK FORECASTING FOR OPTIMAL SATELLITE REMOTE SENSING

3

these can only be measured by the satellite depending on its type of payload(s).

4) In the case of satellite constellations, information sharing between the satellites is critical; each satellite can take advantage of observations by other satellites for the forecasting process. The communication bandwidth between the satellites is, however, limited (16.6–293 kb/s depending on the intersatellite link (ISL) distances for low-cost microsatellites as investigated in [17]), and hence, it is desirable for the shared observation information to be of small size. For example, transmitting the entire raw image captured by a satellite to other satellites in the constellation via ISL may be prohibitive compared with transmitting few observation parameters (e.g., center, velocity, and intensity of TC) derived from the raw image.

5) Must be able to operate for long periods with no intervention from ground stations (i.e., autonomously). Ground-station resources are expensive and sparse. LEO satellites have limited access to ground stations for periods ranging from 5 to 15 min with large gaps between the subsequent revisits, depending on the satellite orbit, number of ground stations, and location of the ground stations. It is desirable to reduce the required amount of data uplinked from the ground stations to the satellites to efficiently utilize the satellite-ground contacts and resources.

The above-outlined requirements motivated the investigation of a methodology and algorithm that can be implemented onboard satellite hardware. While this work concentrates on the development of the algorithm toward a specific natural phenomenon (TCs), the approach of utilizing ensemble member track information for guidance is generic. While the direct contribution of this work is the algorithm geared for forecasting TC tracks, an indirect contribution is the demonstration of the novel utilization of ensemble member track information for the purpose of adaptive forecasting.

III. ALGORITHM

A. Background

The forecast made by NWP models suffers from two types of error: initial conditions errors and inseparable model errors [18]. To account for the uncertainties in the prediction process, ensemble forecasting was established with the ultimate goal to predict quantitatively the probability state of the atmosphere at a future time [18]. A collection of predictions are made for not just the best available estimate of the initial conditions but also for perturbed version of the forecast (by sampling the probability density function of the initial state via Monte Carlo techniques and then evolving those perturbed states using a weather model). The prediction produced from the original forecast is known as the control member, and the predictions from the perturbed versions are known as ensemble members.

The ensemble forecasts represent different possibilities at a later time and hence a set of different, but possible predictions. A simple mean of these possible predictions may be taken to produce the ensemble mean forecast. Taking the concept of the ensembles further, consensus predictions are produced by taking a simple mean of forecasts from different weather models or forming a superensemble [19], which uses past performance of models to determine statistical weights that may vary in both space and time.

The basic idea behind our proposed algorithm is to consider the information from the recent observations (both successful and unsuccessful) to determine which ensemble members are closer to the true state of the dynamical system. The relative importance of the ensemble members is expressed in the form of weights and a linear-weighted average is taken to produce a new forecast track. Extrapolation (after segmented linear regression of raw noisy TC position measurements) of the already observed TC track is carried out to produce a new forecast.

B. Description

The proposed algorithm, termed Ensemble-Guided Cyclone Track Forecast (EGCTF), is outlined in Algorithm 1 along with the definition of the algorithm parameters, inputs, and output. The various parts of the EGCTF algorithm are detailed in the following.

1) Ensemble Tracks (Algorithm 1 Data): The ensemble forecast data considered in this work come from the Global Ensemble Forecast System (GEFS) [20], as obtained from the Automated TC Forecast (ATCF) FTP website [21]. There are 21 track forecasts considered: the control track member "AC00" and the 20 ensemble members (perturbed from the control member) "AP01"–"AP20." An additional "AEMN" mean ensemble track forecast is also present in the GEFS data, which is the mean of all the ensemble members, with all the members weighed equally. The forecasts are issued every 6 h at UTC times 0, 6, 12, and 18 h. The resolution of the forecast time steps is at 6-h time steps, varying from 0 h to a maximum of 384 h. The corresponding "best-track" data [22] (track name denoted as "BEST," also at 6-h rate) is the best estimate of the true path of the TC issued by the NHC, as obtained from the ATCF FTP website. The TC track data used are the forecast TC center positions in latitude and longitude coordinates. The EGCTF algorithm entails the extrapolation of prior estimated TC track using a displacement vector (see Section III-B6), for which uniform measures of distances on a plane are preferred. Therefore, we project the TC track positions onto an XY plane using equidistant conic projection with the two standard parallels placed at equal spacing over the region of TC activity (and thus providing nearly uniform distance measures). Fig. 1 shows the example ensemble and best track plots corresponding to hurricane Florence 2018-09-03 06-h UTC forecast for 24-h period on the XY plane (projection parameters: parallel latitudes at 16.67°N, 33.33°N and central meridian as 45°W). A TC is said to be "captured" successfully when the center of the TC is within the observed area (sensor footprint).

a) Data size: The ensemble track data produced by the NWP model need to be uplinked and synced with all the satellites in a constellation and, therefore, should be of modest size (see Section II). Assuming a 16-bit floating-point

---

Algorithm 1 Ensemble Guided Cyclone Track Forecast (EGCTF)

Data: Ensemble track forecasts from NWP
        Prior ensemble weights
        Prior raw TC center observation data
        Prior failed observations data
Input : Forecast time: t_f
Parameters: Subsegment length: L_{sseg}
    Reward factor: r_f
    Threshold number of prior estimations: n_{etp}^{th}
Output: Forecast TC center at time t_f

1   Process prior raw TC center measurements to get estimated TC center positions. Parameter used here is L_{sseg}.
2   First guess (unnormalized) ensemble weights = Prior ensemble weights.
3   **repeat**
4   Calculate mean ensemble track from ensemble track data and ensemble weights.
5   Calculate *cost*, the average of distances of prior estimated TC centers from mean ensemble track to the set of all estimated TC center positions.
6   Update (unnormalized) ensemble weights according to optimizer strategy.
7   **until** *optimizer convergence*;
8   Reward/ Penalize weights of ensemble tracks according to distance of ensemble track from prior failed observations using the reward factor r_f.
9   Generate *optimal* mean ensemble track from the ensemble weights.
10   **if** *atleast n_{etp}^{th} prior estimates are available* **then**
11   Let the time, position of the immediate prior estimate be $\{X_{est}(t_p), Y_{est}(t_p)\}$.
12   Find TC displacement vector $\overrightarrow{dp}$ using the optical mean ensemble track forecast at t_p and t_f.
13   Extrapolate prior estimate $\{X_{est}(t_p), Y_{est}(t_p)\}$ using the $\overrightarrow{dp}$ to find forecast storm position at t_f.
14   **else**
15   Calculate forecast at t_f as the TC center position from the optimal mean ensemble track only.
16   **end**

Fig. 1. Illustration of successful (sensor footprint on the right) and failed (sensor footprint on the left) capture of TC observations for hurricane Florence 2018, using a sensor with circular footprint of 50-km diameter, as superimposed on the Akima interpolated ensemble tracks and the best track. Note that the scales of the *x*- and *y*-axes are different and that the different points on a track correspond to different times (timestamps shown for one of the ensemble tracks). While two tracks in the 2-D plot may appear to intersect, it may correspond to different times, and hence, in reality, there may be no intersection.

sister applications. For example, the bispectral infrared detection (BIRD) small satellite [23] conducted autonomous feature extraction and classification using a neural network processor. The EO-1 spacecraft has demonstrated onboard generation of the following analysis products: thermal classification and summarization, flood classification, cryosphere classification, quicklook imagery, sulfur detection, and hyperspectral data analysis [24]. Field-programmable gate array (FPGA)-based real-time onboard georeferencing has been proposed in [25].

The geographical position of the center of the TC needs to be estimated from the gathered observation so that this information can be assimilated into our proposed EGCTF algorithm. Estimating TC center is a challenging problem, and the solution is specific to the type of the payload/observation data. Wimmers and Velden [26] described an algorithm to automatically calculate the rotational center of TCs using spirally oriented brightness temperature gradients in the TC banding patterns, along with gradients along the ring-shaped edge of a possible eye. Results are calibrated and validated using 85–92-GHz passive satellite microwave imagery. In [27], the above-mentioned algorithm was improved to consider multisatellite data inputs: 37- and 85–92-GHz microwave imagers; geostationary imagery at visible, near-infrared, and longwave infrared window channels; and scatterometer ambiguities. Jaiswal and Kishtawal [28] demonstrated automatic center determination with image processing techniques applied to satellite infrared images.

In this work, we assume the availability of the TC center position after an observation resulting in an successful TC capture has been made, henceforth called the "raw" storm center measurement, denoted by the coordinate pair $\{X_{raw}(t), Y_{raw}(t)\}$. In a practical scenario, the raw TC center value is not accurate and would possess an associated uncertainty. To improve the estimate of the center position, we consider multiple raw measurements (at different times)

representation of the latitude/longitude values and track data corresponding to a forecast duration of T hours, the amount of data can be calculated as: $T(\text{h}/6 \text{ h}) \times 21$ ensembles $\times 2$ (lat, lon) $\times 16$ (bits). For example, the track data corresponding to 48-h forecast period need a packet size of 672 bytes, without overhead.

2) Processing of Raw TC Center Measurements (Algorithm 1 Line 1): The satellite payload is expected to collect spectral data of the TC. Preprocessing of these data onboard the satellites involves calibration, geometric corrections to enable mapping the image pixels to ground positions, and feature extraction to learn about the current state of the phenomenon. Such technologies have been demonstrated for

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

RAVINDRA *et al.*: ENSEMBLE-GUIDED TC TRACK FORECASTING FOR OPTIMAL SATELLITE REMOTE SENSING 5


Fig. 2. Illustration of the segmented regression method to process the raw TC center measurements to estimated TC centers at the midpoint of the subsegment regression line. Note that while the length (in time dimension) of the subsegments (L_{sseg}) is fixed, the length of the segment is variable and depends on the time sequence of the available raw-measurement.

and form a single "estimated" TC center using the method of segmented regression shown in Fig. 2. A set of observations with no (time) gaps larger than G_{seg} is defined as a curve segment and divided into line subsegments of predefined length L_{sseg} (with $G_{\text{seg}} = L_{\text{sseg}}$). While the parameter L_{sseg} is used to approximate the TC path as a set of line subsegments, the G_{seg} parameter tells the maximum period of absence of data after which a line subsegment cannot be built (and hence, a new curve segment needs to start). Note that the last subsegment is defined with respect to the end of the segment. Within each subsegment of length L_{sseg}, a least mean square linear regression is performed. The midpoint (in time) of the subsegment is defined by the TC center, as estimated from the set of raw TC center measurements of the corresponding subsegment. In case only one raw measurement is available for a subsegment, this measurement is considered to be the estimated TC center. The selection of the parameter L_{sseg} depends on the time over which the TC progress is expected to be linear, which can either be objectively decided based on prior knowledge of the TC path from a study of past similar TCs or by running regional forecast models at a higher temporal resolution or by making a conservative guess, and the noise level of the raw measurements as described in Section IV. The estimated TC center positions at time t are henceforth denoted by the coordinate pair: $\{X_{\text{est}}(t), Y_{\text{est}}(t)\}$.

The abovementioned methodology of segmentation and regression allow for the collation of a dense set of raw TC measurements to obtain a single, more accurate estimation of the TC center. In the case of agile LEO satellites, the pass duration over a small region of interest (such as the region of TC activity) is on the order of 5–15 min, depending on the altitude of the satellite and the maneuverability of the satellite. Several captures are possible over a single satellite pass, which can be collated as described. Furthermore, in the

case of multiple satellites in a single orbital plane (such as in the CYGNSS mission [1]), several successive observations are made by the system of satellites, which can be collated to more accurate estimates.

3) Mean Ensemble Track Calculation [Algorithm 1 *(Line 4)]:* The mean ensemble track is simply the linear-weighted average of all the ensemble track forecasts

$$\{\bar{X}_{\text{ens}}(t), \bar{Y}_{\text{ens}}(t)\} = \{\Sigma_i w^i X_{\text{ens}}^i(t), \Sigma_i w^i Y_{\text{ens}}^i(t)\} \quad (1)$$

where $\{\bar{X}_{\text{ens}}(t), \bar{Y}_{\text{ens}}(t)\}$ are the Cartesian coordinates of the mean ensemble track at time t, $\{X_{\text{ens}}^i(t), Y_{\text{ens}}^i(t)\}$ are the coordinates of the ith ensemble track at time t, and w_i is the weight associated with the ith ensemble track such that $\Sigma_i w_i = 1$ and $w_i \geq 0$.

Since the ensemble forecast data are available at 6-h intervals, the Akima interpolation scheme [29] is applied to obtain TC positions at intermediate times. The Akima interpolation performs piecewise interpolation using third-order polynomials and yields smooth, natural curves through a set of data points. It is used by the NHC to interpolate 12-h best-track data to 6-h intervals [21, NRL_doc_ATCFdatabase.html] [30]. Furthermore, Jarvinen *et al.* [31] reported that the Akima polynomial interpolation gave highly satisfactory results when applied for best-track interpolation from 12- to 6-h intervals, thus making it a favorable interpolation method for the EGCTF algorithm.

4) Optimization of the Ensemble Weights From Set of Estimated Past TC Centers [Algorithm 1 *(Line 3–7)]:* The core of the EGCTF algorithm is the optimization performed to yield a set of optimal ensemble track weights under the constraints of partial information regarding the TC track. The mean ensemble given in (1) is a linear combination of the ensemble tracks. In practice, some ensemble forecasts would be closer to the truth, and by assigning a higher weight to those ensemble forecasts, we can get the corresponding mean ensemble forecast as a better overall forecast.

At the beginning of the forecast period, all the ensemble members are given equal weights and the mean ensemble track calculated would be the same as that of the "AEMN" track in the GEFS forecast data. As observations are made around the area of the TC, information about the presence/absence (identification) of the TC, along with measured (noisy) TC center positions (in case of successful TC captures), is utilized to determine which ensemble members are the better forecasts. The "true" optimal ensemble track can be defined as the linear-weighted average of the ensemble tracks that are closest to the true track (calculation of the true optimal ensemble track would require knowledge of the true track at all times) in terms of the least mean of distances over all times [see the cost function in (2)]. The goal of the optimization portion of the EGCTF algorithm is to converge to the "true" optimal ensemble track using the set of the partial and noisy true track position estimates.

Mathematically, this is framed as an optimization problem with the following attributes.

a) Optimization parameters: The optimization parameters are the unnormalized ensemble member weights $\{w_{\text{ens}}^1, w_{\text{ens}}^2, \ldots, w_{\text{ens}}^{21}\}$. The use of unnormalized weights as

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

6   IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

the parameters avoids imposition of the normalization constraint, i.e., $\{\sum_i w_{\text{ens}}^i = 1\}$ on the optimizer.

b) Cost function: The cost function utilizes the history of all the successful TC captures (and hence the raw TC center measurements) as follows:

$$\text{cost} = \frac{1}{J}\sum_{\forall j} \sqrt{(\bar{X}_{\text{ens}}(t_j) - X_{\text{est}}(t_j))^2 + (\bar{Y}_{\text{ens}}(t_j) - Y_{\text{est}}(t_j))^2} \quad (2)$$

where cost is the average of the distances between the mean ensemble track and the estimated prior TC center positions, j corresponds to the jth prior TC estimation, and J is the total number of prior estimated storm center positions.

c) Optimizer method: The optimizer used is the limited memory Broyden–Fletcher–Goldfarb–Shanno with Boundaries (L-BFGS-B) algorithm [32], a popular quasi-Newton method that allows the user to specify the amount of storage required by L-BFGS-B using a parameter to determine the number of BFGS corrections saved. It also allows for the setting of simple box bounds on some or all of the optimization parameters. The quasi-Newton property and the limited memory property make this optimization method suitable for real-time implementation onboard satellites. The following lower and upper bounds are set for all the optimization parameters (ensemble weights) $0.005\frac{1}{N_{\text{ens}}} \leq w_{\text{ens}}^i \leq 1$, where N_{ens} is the total number of ensemble members. A lower bound of zero is avoided and is set to 0.5% for which the reason is explained in Section III-B5. Since the result of the optimizer is the set of optimal unnormalized ensemble weights, a final normalization step is performed after the optimizer has converged as follows:

$$w_{\text{ens}}^k \longleftarrow w_{\text{ens}}^k / \sum_i w_{\text{ens}}^i \quad (3)$$

where $\longleftarrow$ denotes the update operation.

5) Reward/Penalize Ensemble Tracks Based on Failed Observations [Algorithm 1 (Lines 8 and 9)]: We also use the set of failed observations in the process of the finding the optimal mean ensemble track. In the event, the TC is not captured by a sensor observation, the ensemble tracks further away from the observation position are rewarded by a (reward) factor r_f (units: [per unit distance]), thus effectively penalizing the ensembles close to the failed observation point. This is carried out as follows:

$$x = ||\mathbf{p}_{\text{ens}}^{\mathbf{i}}(t) - \mathbf{p}_{\text{obs}}(t)|| \quad (4)$$
$$w_{\text{ens}}^i \longleftarrow \exp{(r_f x)}w_{\text{ens}}^i \quad (5)$$

where $||\ldots||$ is the vector norm, $\mathbf{p}_{\text{ens}}^{\mathbf{i}}(t)$ is the ith ensemble track forecast position vector, and $\mathbf{p}_{\text{obs}}(t)$ is the observation position vector of the failed observation at time t. This step is carried out for the set of all failed observations and is followed by normalization of the set of ensemble weights. The ensemble mean track constructed from these weights is here referred to as the optimal mean ensemble track.

Equation (5) indicates multiplication with the prior ensemble weight. In a scenario where the prior weight is 0, the reward functionality fails. Hence, it is desired that the ensemble weights from the optimization step [see (3)] have a nonzero
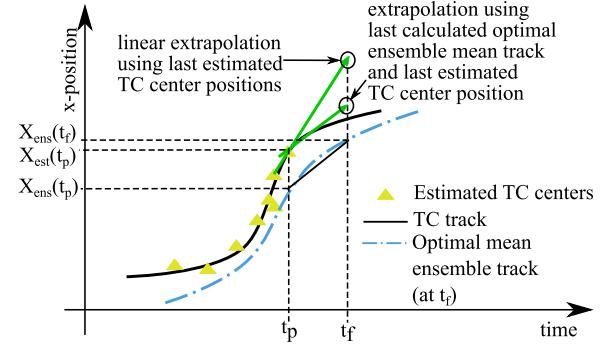

Fig. 3. Illustration of the TC track extrapolation procedure using estimated TC centers at previous timestamps and the optimal mean ensemble track. Those from linear extrapolation with the results are compared. For simplicity, the time trajectory of only the X-coordinate of the TC center is shown.

value. Therefore, the selection of a small positive lower bound in the optimizer variables is described in Section III-B4.

6) Extrapolation of Estimated TC Track [Algorithm 1 (Lines 10–16)]: In an ideal scenario, a large number of ensemble track forecasts would be available such that at least one may potentially be close to the true track. Applying the above ensemble weight optimization would result in a weight ≈ 1 for that ensemble track, whereas for the other ensemble tracks, their weights would be ≈ 0. Producing a large number of ensembles places computational load on the parent NWP model and also on the optimization algorithm since the number of optimization parameters increases. Practically, since the number of available ensemble tracks is limited, the true track may not be expressed as a linear combination of the ensemble tracks. Thus, determining an optimal mean ensemble track by itself is not sufficient for forecasting.

When there are available TC center estimations, one approach is to simply extrapolate linearly the estimated TC track to the forecast time. However, since the TC track is unlikely to be linear, the linear extrapolation would lead to errors whose magnitude depends on: 1) extrapolation time and 2) curvature of the TC track. Therefore, instead of a linear extrapolation, we extrapolate the immediate prior TC center estimate ($\{X_{\text{est}}(t_p), Y_{\text{est}}(t_p)\}$) using the displacement vector derived from the optimal mean ensemble track as follows:

$$\overrightarrow{dp} = \{\bar{X}_{\text{ens}}(t_f) - \bar{X}_{\text{ens}}(t_p), \bar{Y}_{\text{ens}}(t_f) - \bar{Y}_{\text{ens}}(t_p)\} \quad (6)$$

where t_p is the time stamp corresponding to the immediate prior TC center estimate, t_f is the forecast time, and $\overrightarrow{dp}$ is the displacement vector.

Finally, the forecast is calculated by the following vector addition:

$$\{X_f(t_f), Y_f(t_f)\} = \{X_{\text{est}}(t_p), Y_{\text{est}}(t_p)\} + \overrightarrow{dp}. \quad (7)$$

Fig. 3 shows the abovementioned concept. The extrapolation is carried out subject to the condition that a minimum number of prior estimated positions ($n_{\text{etp}}^{\text{th}}$) are available (see Algorithm 1, Line 10). This condition ensures that the corresponding optimal mean ensemble remains close to the true optimal mean ensemble track. In the absence of a minimum number of prior estimated positions, the forecast at time t_f is calculated from interpolation of the optimal mean ensemble track.
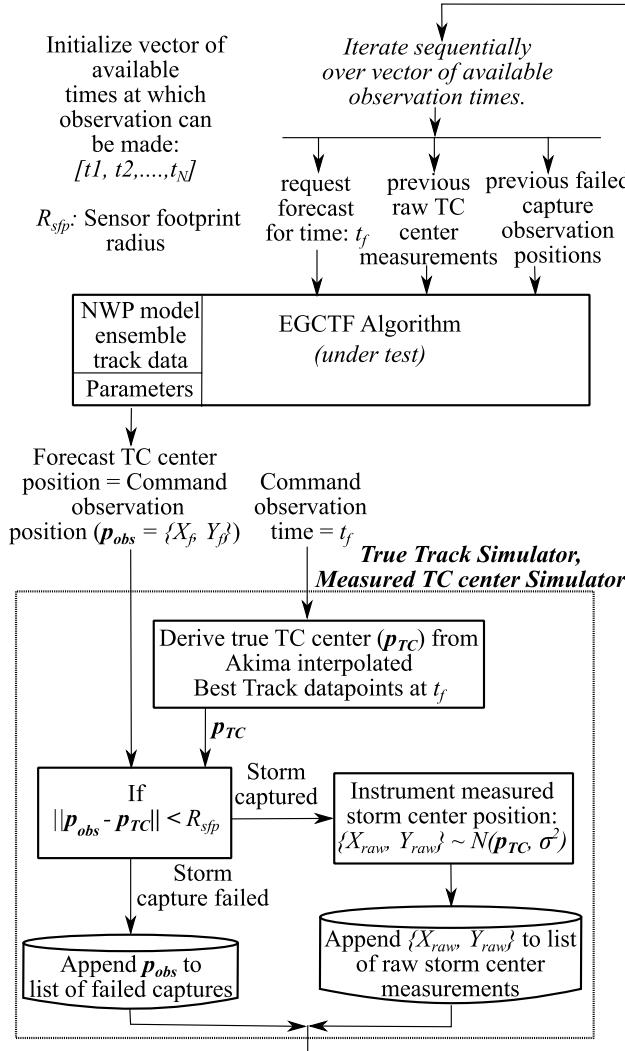
Fig. 4. Testing framework of the proposed EGCTF algorithm (see Algorithm 1). The algorithm under test is placed in loop with the "True Track Simulator" and "Measured TC center Simulator." The test is iterated sequentially over the vector of observation time steps, and the EGCTF algorithm issues forecasts (observation points) for future times while assimilating the results from all previous observations.

IV. TESTS AND RESULTS

A. Framework

The testing framework of the proposed EGCTF algorithm is shown in Fig. 4. The algorithm is initialized with the NWP model ensemble track information corresponding to the TC of interest and parameters that remain constant for a test case. A vector of observation times by a coordinated sensor network (such as a constellation of satellites) is assumed and iterated upon. It should be noted that the forecast at a time t_n is issued immediately after the observation at time t_{n-1} has been made and assimilated by the EGCTF algorithm. The sensor footprint is assumed to be circular with radius R_{sfp}.

The output of the EGCTF algorithm is the predicted storm center position at the requested forecast time that is set to equal the command observation position ($\mathbf{p}_{obs} = \{X_f(t_f), Y_f(t_f)\}$) by the satellite. The "True Track Simulator, Measured TC center Simulator" block uses the 6-h best track data (obtained from NHC [21]), interpolates it using the Akima method, and

treats it as the true track. The position of the TC at time t_f given as $\mathbf{p}_{TC}$ is compared with $\mathbf{p}_{obs}$. If $\mathbf{p}_{TC}$ falls within the sensor footprint, the storm is said to be captured successfully and a raw storm center measurement is calculated as a sample from a bivariate normal distribution with mean $\mathbf{p}_{TC}$ and a covariance σ^2. For a circular sensor footprint of radius R_{sfp}, this condition can be written as

$$||\mathbf{p}_{obs} - \mathbf{p}_{TC}|| < R_{sfp} \implies \{X_{raw}, Y_{raw}\} \sim N(\mathbf{p}_{TC}, \sigma^2).$$

This raw measurement is appended to the list containing previously recorded raw measurements. On the other hand, if the storm capture fails, $\mathbf{p}_{obs}$ is appended to the list containing the previously recorded failed storm capture observation positions. The updated lists are used for the subsequent iteration of the simulation.

B. Test Metrics

There are three metrics used to evaluate the predictive effectiveness of the algorithm.

1) Distance Error Between the Forecast TC Centers (Observation Points) and the Actual TC Center at the Time of Observation: The distance from the forecast TC center to the actual TC center gives an absolute measure of the algorithms' performance. This metric is similar to the track forecast error term used by the NHC [11, p. 4], the difference being that the NHC forecast error utilizes great circle distances and distance error here is measured in the 2-D projection plane in which the EGCTF algorithm operates. While it may be argued that considering the great circle distances is a better objective measure, distances in the projection plane are appropriate for this work since the EGCTF algorithm, and hence the optimization involving minimization of the distance errors (cost), operates in this plane, we would like to test the optimization trend with the defined cost term and are mainly concerned with the relative performance with respect to a baseline forecast model, and both are computed in the same frame.

The distance error (d_e) metric is defined as

$$d_e = ||\mathbf{p}_{obs} - \mathbf{p}_{TC}||. \tag{8}$$

A corresponding average distance error ($\bar{d}_e$) over all the observations made during the course of the simulation is defined as follows:

$$\bar{d}_e = 1/N_{obs} \sum_{N_{obs}} d_e \tag{9}$$

where N_{obs} is the total number of observations over the course of the simulation.

2) Skill With Respect to a Baseline Forecast Algorithm: In keeping with the forecast skill metric defined by the NHC in [11, p. 4], a skill term s is defined as follows to describe the percentage improvement of the EGCTF forecasts over the baseline forecasts:

$$s = 100 \frac{(\bar{d}_e|_{base} - \bar{d}_e|_{EGCTF})}{\bar{d}_e|_{base}} \tag{10}$$

where $\bar{d}_e|_{base}$ is the average distance error yielded by the baseline forecast algorithm and $\bar{d}_e|_{EGCTF}$ is the average distance

error yielded by the EGCTF algorithm for the same simulation scenario. Note that for $d_e|_{EGCTF} < d_e|_{base}$, s is positive, and when $\bar{d}_e|_{EGCTF} = 0$, $s = 100\%$.

The baseline TC forecast referenced here is the AEMN track from the GEFS forecast. The AEMN track is the mean ensemble track with all the ensembles weighted equally. In the absence of the EGCTF algorithm and data assimilation from observations, one may determine the forecasts (and hence the observation positions) from the AEMN track.

3) Total Number of Successful TC Captures at the End of the Simulation Period: The error metrics defined earlier do not differentiate between successful observations (i.e., observations in which the TC center is captured) and failed observations. Hence, we consider a third metric—defined simply as the count of successful observations of TC captures over the entire simulation duration.

C. Test Data and Parameters

1) Track Data: As described in Section III-B, the 21 ensemble members of the GEFS forecasts and the best-track data from NHC were utilized for this work. All TC activity over the Atlantic basin in the year 2018 (16 hurricanes from Alberto to Oscar) was considered. The test simulation period is set to 48 h, and hence, the first 48 h of GEFS forecast track data was used. Some forecasts were disregarded because they did not extend to 48 h and/or some ensemble members were missing. In total, 318 track forecast data and corresponding best-track data were collected. Of these, five were selected to illustrate the behavior of the algorithm at different conditions (observation frequencies and sensor noise). The name of the hurricane and the corresponding forecast epoch time for each of the selected five tracks is listed follows.

TF1: Hurricane Alberto 2018-05-27 06-h UTC.
TF2: Hurricane Florence 2018-09-03 12-h UTC.
TF3: Hurricane Florence 2018-09-08 18-h UTC.
TF4: Hurricane Kirk 2018-09-26 12-h UTC.
TF5: Hurricane Leslie 2018-10-05 18-h UTC.

2) Nominal Values: The nominal values for the algorithm parameters for all the test cases are as follows: $L_{sseg} = 1$ h, $r_f = 1e^{-6}$/m and $n_{etp}^{th} = 5$. The nominal sensor footprint radius was considered to be circular with a diameter of 100 km. The typical diameter of the TC-eye ranges from 30 to 65 km, and the 100-km value was chosen to allow enough buffer to observe sufficient features needed for detecting the TC center. Since all TCs considered were in the Atlantic basin, the following conic equidistant projection parameters are used: parallel latitudes at 16.67°N, 33.33°N and central meridian as 45°W. In practice, when implementing the algorithm for a single TC, the projection parameters may be chosen to minimize the distortions over the region of that particular TC activity. The simulation period is set to 48 h, and the nominal TC center measurement noise covariance matrix is assumed to be diagonal with entries $(10 \text{ km})^2$.

3) Observation Timing Vectors: These vectors contain the times at which the observations are to be made. The test procedures are performed for two types of observation timing vectors.
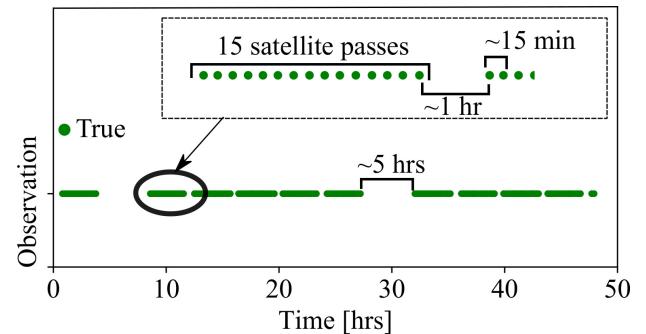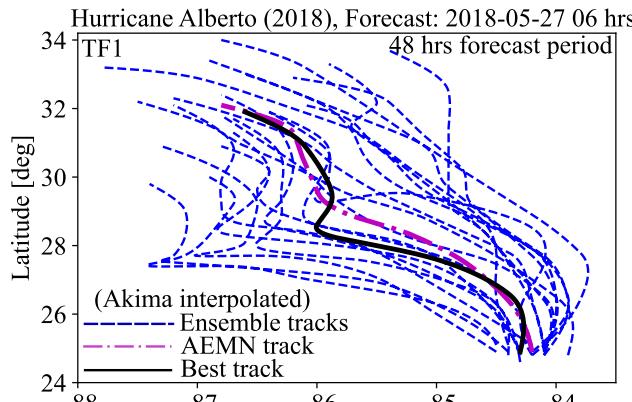

Fig. 5. Simulated observation times available to a Walker-type constellation of satellites (three planes, eight satellites per plane, and orbital plane inclination = 30°), with varying gaps of 15 min, 1 h, and 5 h. One observation is assumed to be made at each satellite pass over the region of TC activity.

1) *Observation Times at Uniformly Spaced but Varying Gap Periods:* These are defined in terms of observation frequency, starting with values of 2/h up-to 50/h in increments of 2/h. The first observation starts at 0.5 h from the ensemble track epoch to account for a possible delay in syncing the sensor network with the NWP model-generated ensemble track forecast.

2) *Observation Times Derived From Orbital Simulation of a Constellation of LEO Satellites Imaging Over the Atlantic Basin:* We considered a Walker constellation [33] with three planes and eight satellites per plane, an altitude of 550 km, and an orbital inclination of 30°. The satellites' field of regards are assumed to be 120° (i.e., the satellite has potential access inside a conical region of (full) cone angle 120°, within which observations can be made). A single observation is deemed to be made during each satellite pass over the region of TC activity. Fig. 5 shows the observation times available to such a constellation—clustered with the variable spacing of about 15 min, 1 h, and 5 h. A cluster of observation times corresponds to a set of 15 observations with roughly 15-min spacing, indicating continuous observation by satellites in two of the three orbital planes. There are a few cases during which more than one satellite has access to the region at the same time, and we simplify such cases by assuming only one effective observation. A total of 161 observation times are present over a 48-h window.
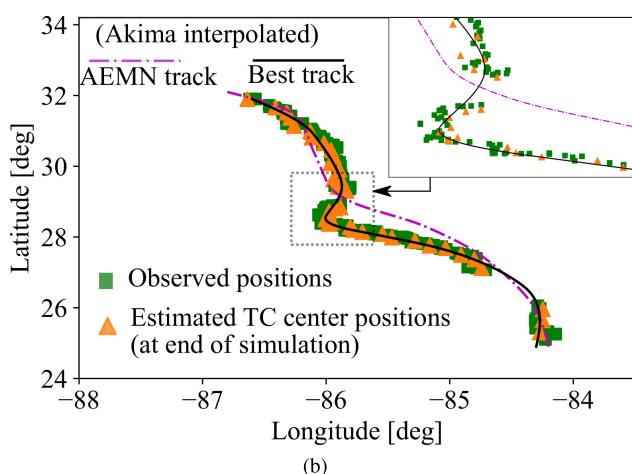
D. Results

The test simulations were run for different algorithm parameters, observation frequencies, and sensor noise levels. The results of testing are presented in Sections IV-D1–IV-D7, starting with a nominal test and its sensitivity to a few critical parameters (1–3), followed by overall behavior of test metrics and the effect of various algorithmic nuances on these metrics (4–7).
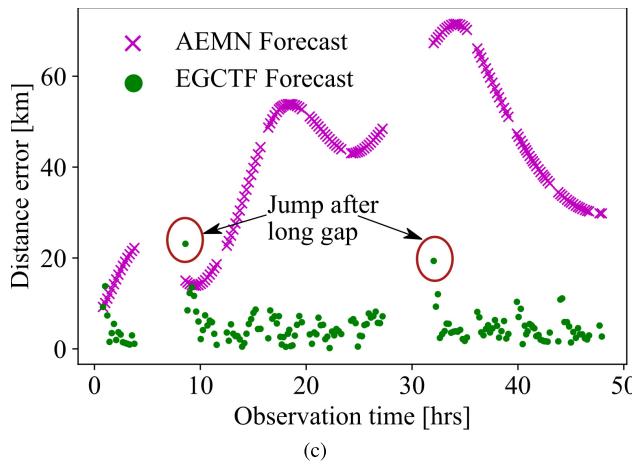
1) Complete Result Description for One Forecast: In this test scenario, the GEFS ensemble forecast of Hurricane Alberto at 2018-05-27 06 h (TF1) is run with the algorithm parameters at their nominal values (see Section IV-C).

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

RAVINDRA *et al.*: ENSEMBLE-GUIDED TC TRACK FORECASTING FOR OPTIMAL SATELLITE REMOTE SENSING 9


(a)

(b)

(c)

Fig. 6. Simulation results for the GEFS ensemble forecasts of Hurricane Alberto at 2018-05-27 06 h (TF1). The observation timing vector corresponds to access opportunities derived by simulating an LEO Walker satellite constellation. The EGCTF algorithm parameters and the simulation parameters are set to the nominal values (see Section IV-C). Application of the EGCTF algorithm resulted in a skill of 88.60%; 100% of TC centers were captured compared with 69% by the baseline method. (a) Akima interpolated ensemble and best (true) tracks. (b) Illustration of the observed positions and the estimated positions with respect to the best track and the AEMN track. All the observations yielded successful TC center captures. (c) Comparison of the d_e error metric for the baseline AEMN forecast and the EGCTF forecast.

The observation times are that of the simulated LEO satellite constellation (see Section IV-C). The selected ensemble forecast corresponds to the one issued for Hurricane Alberto at
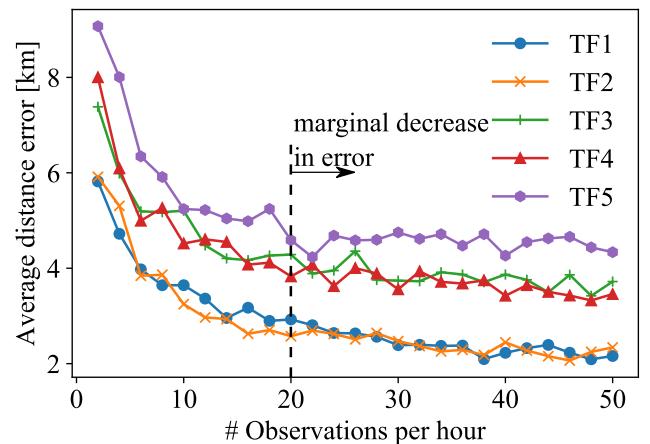

Fig. 7. Average distance error $\bar{d}_e$ as a function of observation frequency for parameters at the nominal values for the five-track forecasts (TF1–TF5) (see Section IV-C). The error is dependent on the track forecast and decreases with an increase in observation frequency. At frequencies higher than 20 observations per hour, the decrease in the error is seen to be marginal, signaling diminishing returns.

2018-05-27 06 h. While the AEMN forecast resulted in $\bar{d}_e = 41.4$ km and 111 successful observations (69%), the EGCTF forecasting method resulted in $\bar{d}_e = 5.4$ km and 161 successful observations (100%) and, therefore, a skill of $s = 88.60\%$. Fig. 6 shows the results of the simulation. Both the observed positions and the estimated TC center positions are shown for the entire time period. The effect of segmented regression in producing more accurate estimates of the TC center positions is illustrated in the zoomed-in inner figure in Fig. 6(b)—the observation positions (green) yielded by the EGCTF algorithm can be seen centered around the best track (black). The observation positions yielded by using the AEMN track as the forecast method (not shown in the figure for purpose of clarity) would be along the purple AEMN track and can be seen to be relatively inaccurate. The effect of the long observation gap after 5 and 28 h can be seen in the form of large d_e for the observations made after these gaps. However, the large error is short-lived and the information from each of these observations is used to correct the subsequent observations, leading to lower error thereafter.

Appendix V contains the results for the remaining track forecast cases TF2–TF5, with the parameters set at nominal values, and for the observation times corresponding to a constellation of LEO satellites.

2) Sensitivity to Observation Frequencies: In this test scenario, the forecast/observation error $\bar{d}_e$ at different observation frequencies (observation times with uniform gaps) was investigated for the test cases of TF1–TF5, and the results are shown in Fig. 7. All the parameters are set to nominal values, as described in Section IV-C, and the observation frequency is varied from 2 to 50 observations per hour in increments of two observations per hour. As expected, the forecast error exhibits a general downward trend as the observation frequency increases since more frequent assimilation takes place. While the decrease in the forecast error is more evident going from two observations per hour to 20 observation per hour, it is marginal at higher observation frequencies. This indicates the

(a)

(b)
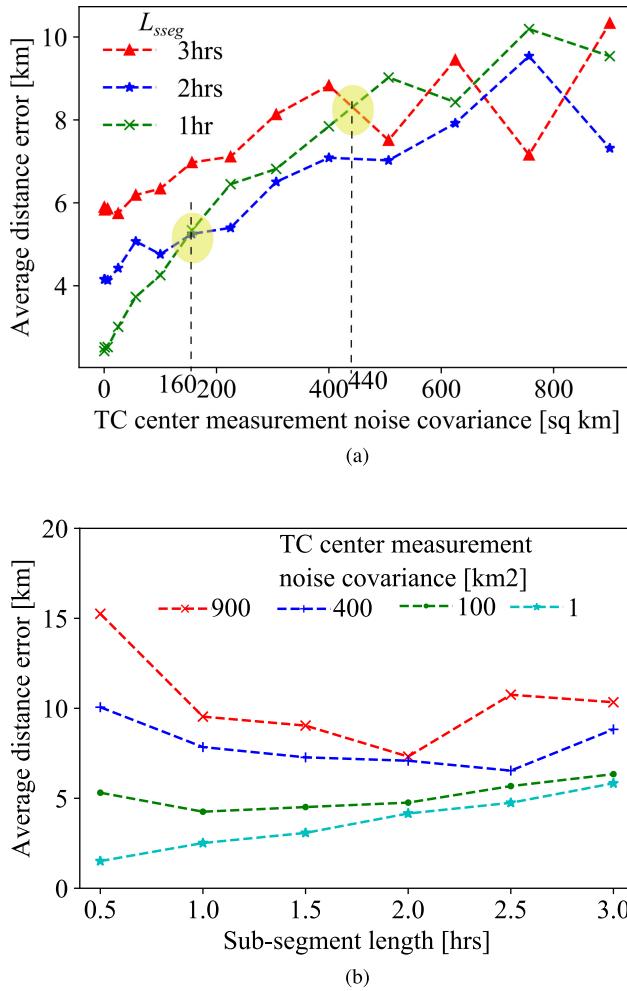
Fig. 8.   Forecast/observation error as a function of TC center measurement noise covariance for test simulation TF1. All other nominal parameters are used. The observation vector was set to a uniform observation rate of 5/h. Error increases with noise as expected. The behavior with respect to variation in L_{sseg} depends on the noise level. For small noise covariance, increasing L_{sseg} results in increase in the error, whereas for large noise covariance, increasing L_{sseg} decreases the error. (a) $\bar{d}_e$ as a function of noise covariance of the TC center measurements for three sets of L_{sseg}. (b) $\bar{d}_e$ as a function of L_{sseg} for four sets of noise covariances.

limitation of the EGCTF algorithm, which uses only measures of the observed TC center positions for the assimilation and does not consider other TC attributes. Furthermore, for a given observation frequency, different tracks exhibit different errors, which can be attributed to the quality of the ensemble track forecasts. By visually inspecting the ensemble plots and best track plots of TF1 shown in Fig. 6 and of TF2–TF5 in Fig. 15, it is possible to get a sense of the relative quality of the ensemble forecasts and map it back to the relative errors between the different track forecast simulations in Fig. 7.

3) Sensitivity to TC Center Measurement Noise Covariance: In this test scenario, we investigate the performance of the EGCTF algorithm for different magnitudes of noise covariance of the TC center measurement (raw observations) and subsegment length L_{sseg}. Fig. 8 shows the results for the TF1 scenario where the EGCTF algorithm is run with the set of nominal parameters (variables being the noise covariance and L_{sseg}). A uniform observation frequency of five observations
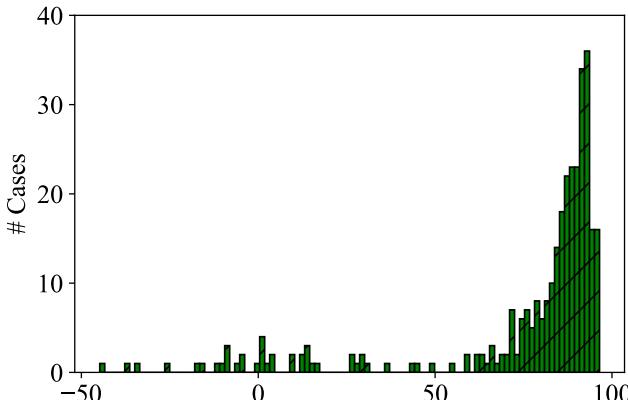
per hour is chosen to be representative of observations by an imminently implementable satellite constellation. The noise covariance was varied from $(0.5 \text{ km})^2$ to $(30 \text{ km})^2$, while L_{sseg} was varied from 0.5 to 3 h. With the value of L_{sseg} as 0.5 h (and hence $G_{\text{seg}} = 0.5$ h), and the observation frequency as five observations per hour, note that the tests involved building one continuous curve segment in all the cases (see Fig. 2). As expected, the error increases with increase in noise magnitude (and hence decreasing quality of the observations being assimilated [see Fig. 8(a)]).

The behavior with respect to L_{sseg} on the other hand exhibits a strong dependence on the underlying noise covariance considered in the simulation, per Fig. 8(b). The error increases with increasing L_{sseg} for small noise covariance, whereas the error is seen to decrease for large noise covariance. With the observation frequency fixed, the number of observations that are collated to form the single TC center estimate depends on the length of subsegment L_{sseg}. While a larger number of collated observations from large L_{sseg} allow for a more accurate estimate, there are three drawbacks on using large L_{sseg}, which increases error.
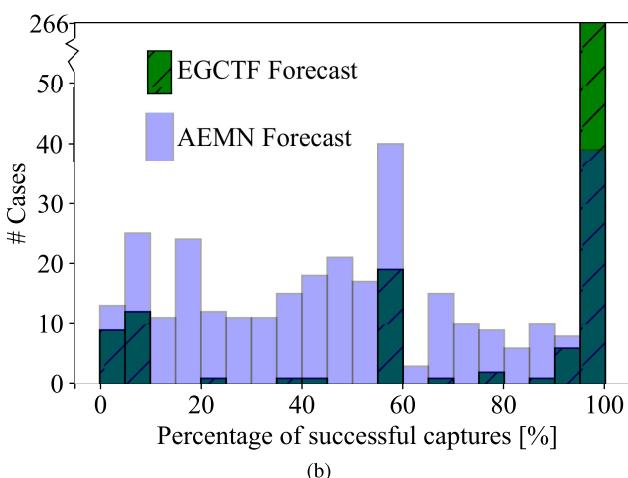
1) The TC may not take a linear path over the L_{sseg} length of time, and in general, the linear segment approximation of the TC path loses accuracy with increasing lengths.
2) There is a reduction of the number of estimates made and the estimates are further spaced apart from each other, thus effectively lowering the resolution at which the prior track is estimated.
3) Since the extrapolation is carried out from the immediate prior estimate which is at the center of the subsegment to the forecast time, the extrapolation is made over a longer time for increasing subsegment length. A shorter extrapolation time is desired for more accurate forecasts.

Furthermore, from Fig. 8(a), the following observations can be made comparing the three colored plot lines corresponding to different values of L_{sseg}. At lower noise conditions ($<160 \text{ km}^2$), there is little benefit of using larger subsegment lengths (L_{sseg}) for estimation. The drawbacks of using larger L_{sseg} outweigh the benefits leading to larger errors. The error plot lines cross each other at the circles marked in yellow, indicating that for noise covariances higher than those, a higher L_{sseg} exhibits lower error and is hence desirable. For example, for noise covariances greater than 160 km^2, L_{sseg} of 2 h exhibits lower error than that of L_{sseg} of 1 h.

4) Average Behavior Over All Forecasts: In this scenario, the forecast error for all 318 track forecasts (GEFS ensemble members and NHC best tracks) for the 2018 year hurricanes in the Atlantic basin was evaluated. The relative performance with respect to the chosen baseline method (using forecasts from the AEMN track) is presented. All the parameters are at nominal values and the observation times are derived from the LEO satellite constellation described in Section IV-C. Fig. 9 shows the histogram of the skill [defined in (10)] and comparison of the percentage of successful TC captures made using the EGCTF versus AEMN forecasts. Of the 318 track forecasts and test cases, the EGCTF algorithm exhibits better performance in terms of the defined skill for 303 cases. In 247

(a)
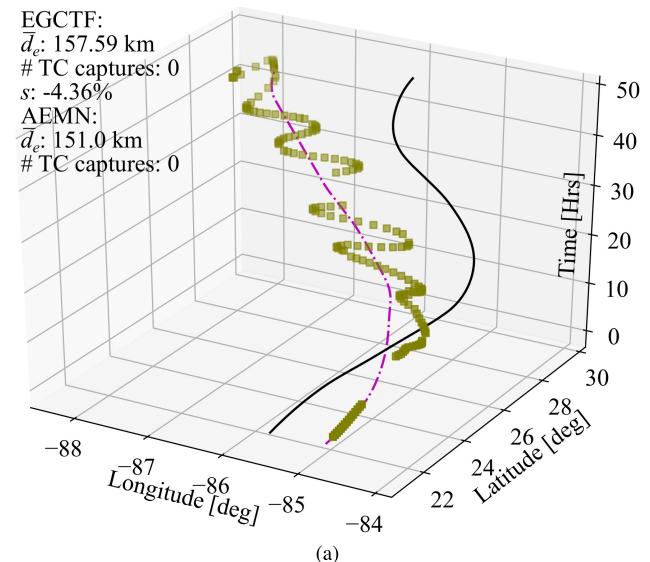
(b)
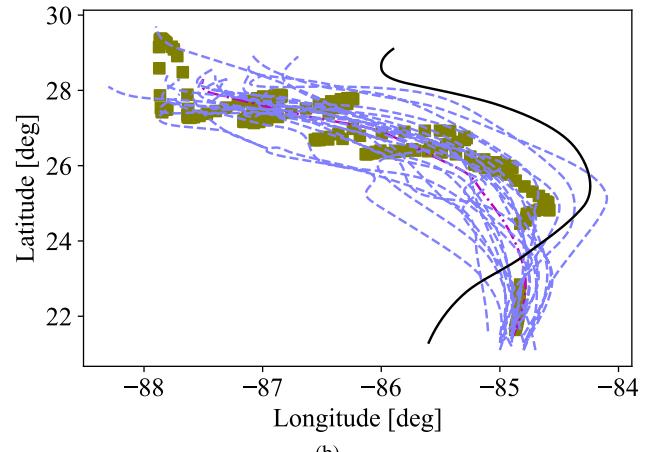
Fig. 9. Histograms illustrating the results of the 318 test cases utilizing the 318 GEFS ensemble track forecasts of the 2018 hurricane season in the Atlantic basin. Nominal set of parameters was used, and the observation times were derived from the observation opportunities available to an LEO satellite constellation (see Section IV-C). The EGCTF exhibits better skill and a larger number of successful TC captures, compared to using AEMN forecasts. (a) Histogram of forecast skill of the EGCTF algorithm, i.e., improvement over the baseline forecast from AEMN tracks. (b) Comparison of the number of successful TC captures over the entire simulation period of 48 h using the EGCTF versus AEMN forecasts.

cases (over three quarters of the total number of test cases), the performance of the EGCTF algorithm exhibits skill greater than 75%. In a total of 161 observation opportunities available in the observation times vector, the EGCTF algorithm was able to capture the TC in over 90% of the opportunities for 271 test cases (last two green bars). The baseline method could capture over 90% of the opportunities in only 47 test cases (last two purple bars). Furthermore, only in 7 of 318 test cases does the baseline method outperform the EGCTF forecast in terms of the number of successful TC captures.

While Fig. 9 confirms that the EGCTF generally performs better than the baseline, we also investigated the cases where the EGCTF algorithm underperforms. Fig. 10 shows one such test case where the skill was negative ($s = -4.36\%$). It discerned from the ensemble tracks and best track plots that the GEFS ensemble member forecasts are poor predictions of the TC track. The high initial error of the NWP model predictions is visible in the form of the ensemble tracks starting at a

(a)

(b)

Legend:

| (Akima interpolated) | ☐ Observed positions (successful) |
- - - - Ensemble tracks	☐ Observed positions (unsuccessful)
—·—·— AEMN track	▲ Estimated TC center positions
—— Best track	(at end of simulation)
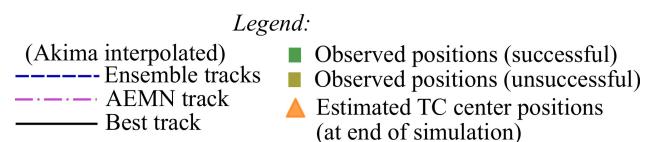
Fig. 10. Illustration of the track, observed positions, and estimated TC center positions of test cases where the EGCTF algorithm underperforms, resulting in no successful captures and no estimated TC center positions. (a) 3-D plot of tracks and observation positions, with time on the vertical axis (ensemble tracks are not included for clarity). (b) 2-D plot of tracks and observation positions.

location relatively far away from the actual TC (compared with the closeness of the ensemble and best tracks in Figs. 6(a) and 15). The sequence of observation positions appears to oscillate around the AEMN track (searching for the true track) because the EGCTF algorithm rewards ensembles predictions that are farther away from the failed observation point. However, the true TC being far out, no successful captures were made during the course of observations, therefore the relatively poor performance.

5) Effect of Utilizing the Optimal Ensemble Track Displacements for Extrapolation: The optimal ensemble track displacements from the time of issue of the forecast to the forecast timestamp are utilized in the extrapolation section of
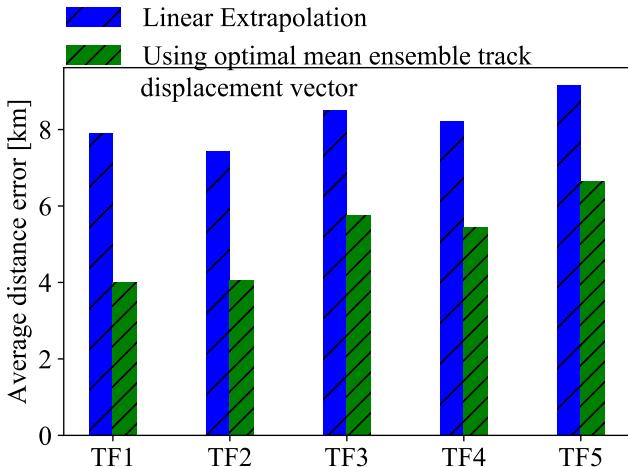
Fig. 11. Comparison of results from EGCTF test runs using displacement vectors from the optimal ensemble track for forecasts using extrapolation versus a modified version of the EGCTF algorithm using linear extrapolation. All the parameters are set to the nominal values (see Section IV-C). The observation timing vector corresponds to a uniform observation rate of five observations per hour. The native EGCTF algorithm resulted in lower forecast errors ($\bar{d}_e$).

the EGCTF algorithm, as described in Section III-B. In this test scenario, we compare the resulting forecast errors of the test conducted with the EGCTF algorithm versus the test conducted with a modified version of the EGCTF algorithm, which uses linear extrapolation of the prior observed TC center positions. The tests were conducted for track forecasts TF1–TF5 using the nominal set of parameters given in Section IV-C. As seen in Fig. 11, the errors of the tests that used the modified version of the EGCTF (with linear extrapolation) are relatively higher (by at least 25%). This demonstrates the effectiveness of using guidance from the optimal ensemble track for extrapolation.

6) Effect of Taking Failed Observations Into Consideration: To show the effect of including information from the failed observations in the form of an exponential rewarding function (which uses the distance between the failed observation point and the ensemble member forecast as input) and the selection of the reward factor r_f parameter, the EGCTF algorithm was run on an example case for different values of r_f, and all other parameters kept at their nominal values (see Section IV-C). The observation times are that of the simulated LEO satellite constellation (see Section IV-C). Fig. 12 shows the results of this test scenario on the track forecasts of hurricane Florence (epoch: 2018-09-12 18 h).

A $r_f = 0$ (lower bound) implies multiplication of the ensemble weights by unity, and hence, effectively, the failed observations are not taken into consideration. In Fig. 12(a), it can be seen that the optimization algorithm converges to the peripheral ensemble track member (marked "A," far from the true track) due to the set of successful TC captures made initially and continues to favor that track for the rest of the simulation period.

An upper bound can be set on r_f depending on the maximum spread of the ensemble set. In an extreme case, a failed observation can happen near the peripheral track of
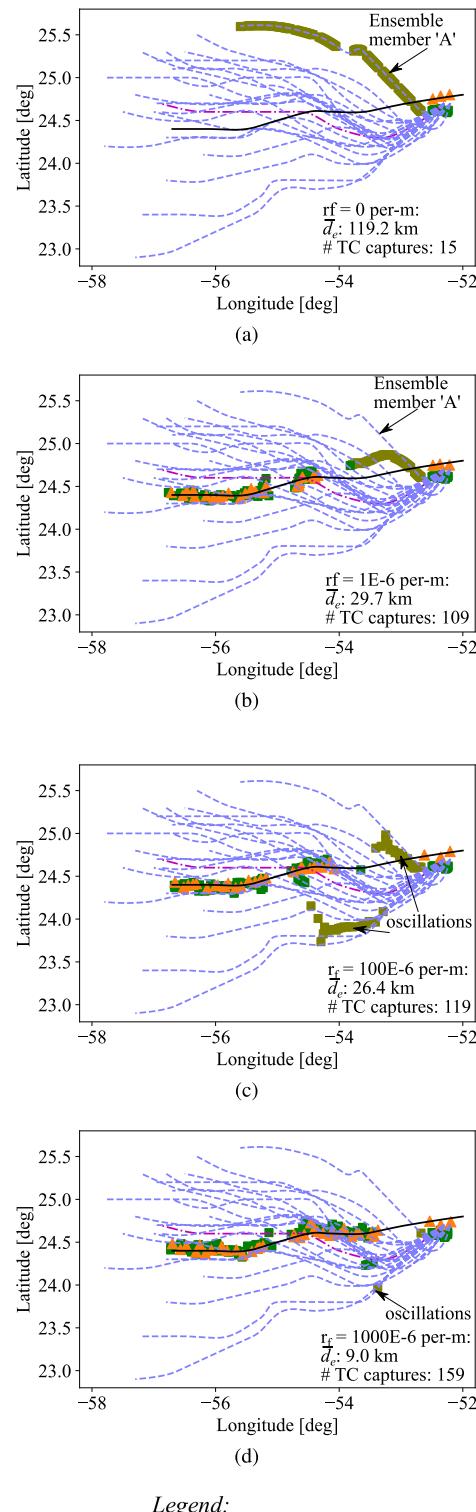
Fig. 12. Comparison of results of test simulations with EGCTF with different values of the reward factor r_f. All other algorithm parameters are set to the nominal values (see Section IV-C). The ensemble track forecasts and the best track corresponds to hurricane Florence at epoch of 2018-09-12 18 h and a simulation duration of 48 h. (a) $r_f = 0E6/m$. (b) $r_f = 1E - 6/m$ (nominal value used in the tests). (c) $r_f = 100E - 6/m$. (d) $r_f = 1000E - 6/m$.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

RAVINDRA *et al.*: ENSEMBLE-GUIDED TC TRACK FORECASTING FOR OPTIMAL SATELLITE REMOTE SENSING   13

an ensemble set. The ensemble track farthest away from the failed observation (at the other end of the periphery) shall be weighted by a large factor: $\exp(r_{\mathrm{fmax}}x_{\mathrm{max}})$, where x_{max} is the maximum expected distance between a failed observation and an ensemble track and r_{fmax} is the maximum reward factor that can be used. The term $\exp(r_{\mathrm{fmax}}x_{\mathrm{max}})$ should be less than the overflow trigger value (depending on the underlying processing system). For example, if the processor uses the IEEE 745 32-bit floating-point representation, then $r_{\mathrm{fmax}} = \log_e(3.4E38)/x_{\mathrm{max}}$. $x_{\mathrm{max}} = 100E3$ m gives $r_{\mathrm{fmax}} = 887.2E - 6$/m. The x_{max} term can be determined from the set of ensemble tracks.
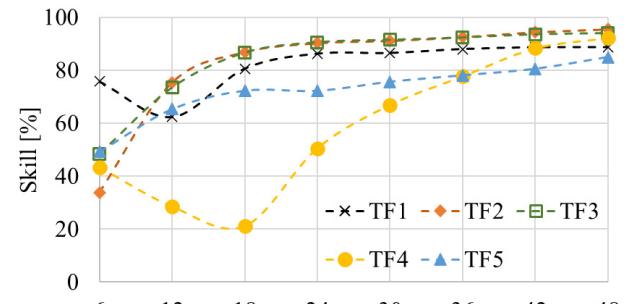
In Fig. 12(b), the rewarding of ensemble members far from the failed observation point steers the course of observations gradually toward the true track. In general, the reward factor causes oscillations about the previous optimal mean ensemble track as can be seen in Fig. 10, where there are no successful observations, and the series of failed observations results in the algorithm searching for the TC about the mean ensemble track. The sensitivity to increasing magnitudes of r_f is shown in Fig. 12. The amplitude and frequency of the oscillations increase with the magnitude of r_f as expected. While a large reward factor ($r_f = 1000E - 6$/m) results in a better performance, the selection of a small r_f is a more conservative choice so that the search is carried out smoothly (without discontinuities and potential to miss the TC center) over the ensemble spread.

7) Variation of Performance Metrics With Forecast Period: The test simulations described earlier considered a total forecast period of 48 h, which would enable autonomous operation of the satellites for that period. Since updates from GEFS are available periodically at a period of 6 h, it is possible to update the satellite network with new ensemble information and, hence, restart the EGCTF algorithm with new ensembles every 6 h, assuming that ground-station resources (contacts and station bandwidth) are available to the satellite network every 6 h.
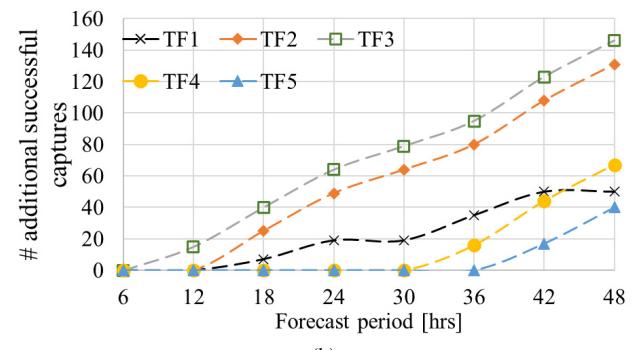
We explored the dependence of the EGCTF algorithm metrics on the forecast period in steps of 6 h up to 48 h, assuming ensemble updates from ground station to the satellite network at those respective intervals. All other simulation parameters are kept at their nominal values and observation times are that of the simulated LEO satellite constellation (see Section IV-C).

Fig. 13 shows the variation in the level of skill and number of additional successful TC captures for different forecast periods for test cases TF1–TF5. The results can be cross-referenced with Figs. 6 and 15, which illustrate the ensemble tracks, AEMN track, and the observation positions for test cases TF1–TF5. Both the metrics are seen to improve, in general, with increasing forecast period because forecasts from the AEMN tracks worsen, while the forecasts from the EGCTF algorithm maintain good performance due to regular assimilation of observations. A positive skill and an equal or more number of successful TC captures were observed over all the forecast periods.

Fig. 14 shows the results of the number of cases exhibiting skill greater than 0%, 25%, 50%, and 75% over different


Fig. 13. EGCTF metrics evaluated for different forecast periods for the cases TF1–TF5. The performance of the algorithm (with respect to the baseline method of using forecasts from the AEMN track) improves for longer forecast periods. (a) Skill as a function of forecast period. Positive skill was observed for all periods. (b) Number of additional successful TC captures using the EGCTF algorithm (over and above the baseline method) as a function of forecast period. In all the cases, the same or more number of successful captures was observed when the EGCTF algorithm was used. The total number of possible captures (observation opportunities) was 161.
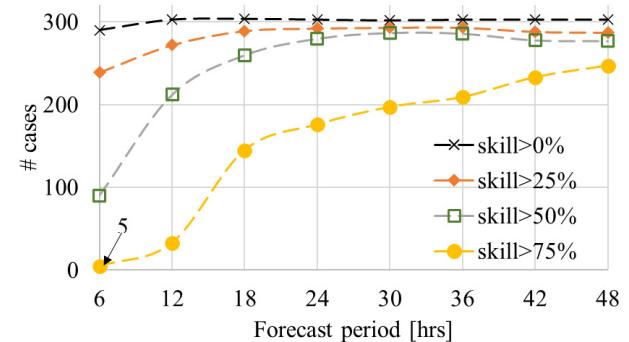

Fig. 14. Number of cases with skill greater than 0%, 25%, 50%, and 75% is shown for different forecast periods for the 318 test cases utilizing the 318 GEFS ensemble track forecasts of the 2018 hurricane season in the Atlantic basin. Nominal set of parameters was used and the observation times were derived from the observation opportunities available to the LEO satellite constellation (see Section IV-C).

forecast periods for the 318 test cases using the 318 GEFS ensemble track forecasts of the 2018 hurricane season in the Atlantic basin (see Section IV-C); 290 cases exhibited positive skill for the forecast period of 6 h and saturated to 303 cases with positive skill at a forecast period of 48 h.

A note of caution must be considered while reading the evaluated skills at the lower forecast periods (less than 12 h).
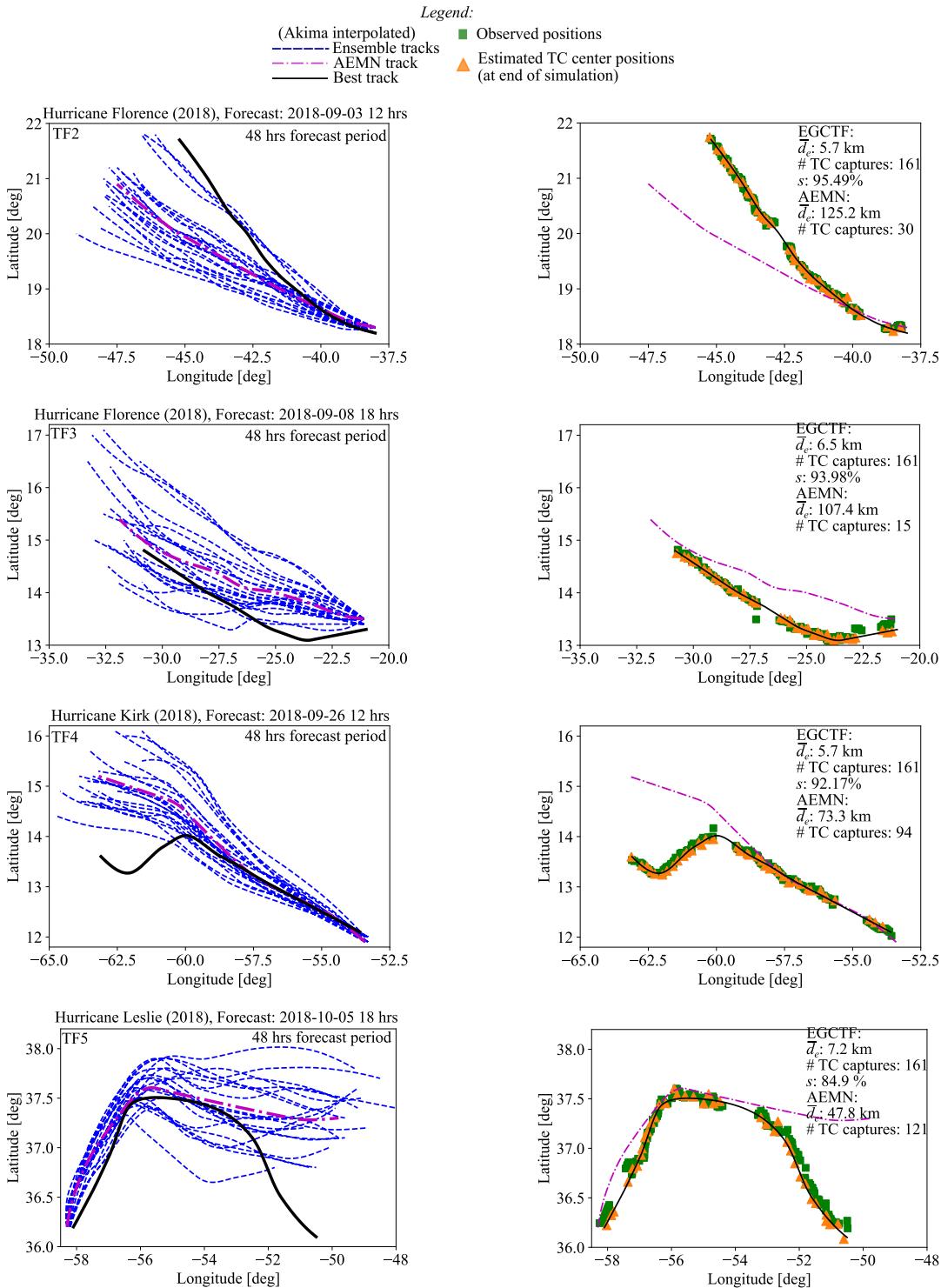
Fig. 15. Ensemble, AEMN, best tracks, and EGCTF simulation results for test cases using track forecasts TF2–TF5. The observation timing vector corresponds to the access opportunities by the constellation of LEO satellites described in Section IV-C. All the observations yielded successful TC center captures.

The skill was seen to be varying significantly over different runs of the experiment, and the presented results are of one such simulation run (a sample). Since the raw measurements are noisy [nominal noise variance of $(10 \text{ km})^2$, the predicted TC center positions are not the same for different experiment runs, although each run does use the same set of simulation parameters and observation times. The effect (variability of result) on skill is prominent during the lower forecast periods as the algorithm is yet to converge to the true track. The effect on the number of observations is relatively low.

V. Conclusion and Future Work

We have proposed and demonstrated an algorithm (EGCTF) that utilizes ensemble track forecasts as guidance while

assimilating TC center measurements to improve previous forecasts on the fly. The low computational load and the use of only the TC center measurements from the payload data make this algorithm a viable option for implementing on a network of interconnected distributed sensors with sparse connectivity, such as a constellation of satellites. The size of the ensemble forecast data, which needs to be uplinked to the satellites for the execution of the algorithm, was calculated to be a modest quantity in the order of a few hundred bytes. Only the observation times, positions, and the TC center estimates (derived from the raw observation data) need to be shared among the network of sensors, which is feasible within commonly available communication bandwidth between the individual satellites.

The proposed algorithm consists of the following main components to predict the TC at a future timestamp: 1) regression of noisy TC center measurements over a predefined (time) length to produce more accurate TC center estimates; 2) use of estimated TC center positions to find optimal ensemble weights; 3) use of failed observation positions to further refine the weights; 4) formulation of the optimal mean ensemble; and 5) extrapolation from the immediate prior TC center estimate to the forecast timestamp using the displacement vector, as calculated from the optimal mean ensemble. The algorithm was tested with data derived from the hurricanes in the year of 2018 within the Atlantic basin. The GEFS 21 ensemble members were used as the ensemble track inputs for the test cases, and the NHC Best Track data were used as the truth data. Access opportunities by an LEO satellite constellation for a region over the Atlantic basin were simulated using orbital dynamics and available opportunities used as the observation times. Uniformly spaced observation times (at different periods) were also explored as comparative observation timing vectors.

Metrics of evaluation include the distance error d_e, average distance error $\bar{d}_e$, skill s, and the number of successful captures. Results for the test scenario with the observation timings derived from the LEO satellite orbital simulation showed positive skill for 303 cases of the total 318 cases simulated for a forecast period of 48 h. The skill was over 75% for 247 test cases, and TC centers were captured more than 90% of the time in 271 test cases. Sensitivity to various aspects of the EGCTF algorithm was explored via the tests, such as the behavior with respect to noise levels of the TC center measurements, subsegment length L_{sseg} (an algorithm parameter), consideration of ensemble track displacement for the extrapolation process, and rewarding tracks based on distance from a failed observation point. In terms of increasing observation frequency, it was found that beyond an observation rate of 20 observations per hour, the algorithm provided diminishing returns in terms of error. Nonetheless, the obtained error ($\bar{d}_e$) was smaller than that of the TC center measurement standard deviation noise. Error also depends on the particular ensemble track forecast, suggesting that better ensemble member forecast data can lead to improved performance. All comparative metrics were seen to improve for longer forecast periods because forecast quality by the baseline method degrades, but the quality of forecast from

the EGCTF algorithm remains steady due to the assimilation of observations.

While in this work we considered ensemble forecasts input from a single NWP model (GEFS), the use of ensemble track forecasts from multiple models such as the European Center for Medium-Range Weather Forecasts Ensemble Prediction System (ECMWF EPS) is of interest as future work. Furthermore, depending on the payload data, it may be possible to obtain other types of TC information from an observation beyond simply its center position, such as the direction of the cyclone and intensity. Incorporating this information into the assimilation process could lead to better forecasts. Another aspect, which remains to be incorporated is a "smooth continuation" when updated ensemble information, is provided from the ground station. The forecasting process can be continued by simply restarting the EGCTF algorithm with the new ensemble information; however, performance may be discontinuous if the ensemble data have initial bias error. A smooth continuation can be explored by using the previous observation and ensemble information. Finally, while this article demonstrates the application of ensemble forecast as guidance for on-the-fly prediction of TC tracks, the same concept is relevant to predicting other natural phenomena with the goal of coordinating satellite observations.

APPENDIX
MORE PLOTS OF THE TRACKS AND OBSERVATION POINTS
USED IN THE TESTS

See Fig. 15.

REFERENCES

[1] C. S. Ruf *et al.*, "A new paradigm in Earth environmental monitoring with the CYGNSS small satellite constellation," *Sci. Rep.*, vol. 8, no. 1, pp. 1–3, Jun. 2018.

[2] Z. Cui, Z. Pu, V. Tallapragada, R. Atlas, and C. S. Ruf, "A preliminary impact study of CYGNSS ocean surface wind speeds on numerical simulations of hurricanes," *Geophys. Res. Lett.*, vol. 46, no. 5, pp. 2984–2992, Mar. 2019.

[3] E. Peral *et al.*, "Raincube, the first spaceborne precipitation radar in a 6U cubesat," in *Proc. 31st Annu. AIAA/USU Conf. Small Satell. (SSC)*, 2019, pp. 1–7.

[4] S. C. Reising *et al.*, "An Earth venture in-space technology demonstration mission for temporal experiment for storms and tropical systems (Tempest)," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. IGARSS*, Jul. 2018, pp. 6301–6303.

[5] W. J. Blackwell *et al.*, "Overview of the NASA TROPICS cubesat constellation mission," *Proc. SPIE*, vol. 10769, Sep. 2018, Art. no. 1076908.

[6] S. Nag, T. Hewagama, G. T. Georgiev, B. Pasquale, S. Aslam, and C. K. Gatebe, "Multispectral snapshot imagers onboard small satellite formations for multi-angular remote sensing," *IEEE Sensors J.*, vol. 17, no. 16, pp. 5252–5268, Aug. 2017.

[7] S. Nag *et al.*, "Autonomous scheduling of agile spacecraft constellations with delay tolerant networking for reactive imaging," in *Proc. Int. Conf. Automated Planning Scheduling SPARK Workshop*, 2019, pp. 1–10.

[8] A. G. Davies *et al.*, "Monitoring active volcanism with the autonomous sciencecraft experiment on EO-1," *Remote Sens. Environ.*, vol. 101, no. 4, pp. 427–446, Apr. 2006.

[9] F. Ip *et al.*, "Flood detection and monitoring with the autonomous sciencecraft experiment onboard EO-1," *Remote Sens. Environ.*, vol. 101, no. 4, pp. 463–481, Apr. 2006.

[10] C. Roy and R. Kovordányi, "Tropical cyclone track forecasting techniques—A review," *Atmos. Res.*, vols. 104–105, pp. 40–69, Feb. 2012.

[11] J. P. Cangialosi, "National hurricane center forecast verification report 2018 hurricane season," Nat. Hurricane Center, Miami, FL, USA, Tech. Rep., Jun. 2019. [Online]. Available: https://www.nhc.noaa.gov/verification/pdfs/Verification_2018.pdf

[12] E. N. Rappaport *et al.*, "Advances and challenges at the national hurricane center," *Weather Forecasting*, vol. 24, no. 2, pp. 395–419, Apr. 2009.

[13] R. Kovordányi and C. Roy, "Cyclone track forecasting based on satellite images using artificial neural networks," *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 6, pp. 513–521, Nov. 2009.

[14] M. Rüttgers, S. Lee, S. Jeon, and D. You, "Prediction of a typhoon track using a generative adversarial network and satellite images," *Sci. Rep.*, vol. 9, no. 1, pp. 1–15, Apr. 2019.

[15] M. M. Kordmahalleh, M. G. Sefidmazgi, and A. Homaifar, "A sparse recurrent neural network for trajectory prediction of atlantic hurricanes," in *Proc. Genetic Evol. Comput. Conf. GECCO*, 2016, pp. 957–964.

[16] *Instruments: Advanced Baseline Imager (ABI)*. Accessed: Nov. 18, 2019. [Online]. Available: https://www.goes-r.gov/spacesegment/abi.html

[17] B. C. Gunter and D. C. Maessen, "Space-based distributed computing using a networked constellation of small satellites," *J. Spacecraft Rockets*, vol. 50, no. 5, pp. 1086–1095, Sep. 2013.

[18] M. Leutbecher and T. N. Palmer, "Ensemble forecasting," *J. Comput. Phys.*, vol. 227, no. 7, pp. 3515–3539, Mar. 2008.

[19] T. N. Krishnamurti, V. Kumar, A. Simon, A. Bhardwaj, T. Ghosh, and R. Ross, "A review of multimodel superensemble forecasting for weather, seasonal climate, and hurricanes," *Rev. Geophys.*, vol. 54, no. 2, pp. 336–377, Jun. 2016.

[20] X. Zhou, Y. Zhu, D. Hou, Y. Luo, J. Peng, and R. Wobus, "Performance of the new NCEP global ensemble forecast system in a parallel experiment," *Weather Forecasting*, vol. 32, no. 5, pp. 1989–2004, Oct. 2017.

[21] *Index of /atcf*. Accessed: Nov. 18, 2019. [Online]. Available: ftp://ftp.nhc.noaa.gov/atcf

[22] C. W. Landsea and J. L. Franklin, "Atlantic hurricane database uncertainty and presentation of a new database format," *Monthly Weather Rev.*, vol. 141, no. 10, pp. 3576–3592, Oct. 2013.

[23] W. Halle, K. Brie, M. Schlicker, W. Skrbek, and H. Venus, "Autonomous onboard classification experiment for the satellite BIRD," *Int. Arch. Photogramm. Remote Sens. And Spatial Inf. Sci.*, vol. 34, no. 1, pp. 63–68, 2002.

[24] S. Chien, D. Mclaren, D. Tran, A. G. Davies, J. Doubleday, and D. Mandl, "Onboard product generation on Earth observing one: A pathfinder for the proposed hyspiri mission intelligent payload module," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 257–264, Apr. 2013.

[25] D. Liu *et al.*, "On-board georeferencing using FPGA-based optimized second-order polynomial equation," *Remote Sens.*, vol. 11, no. 2, p. 124, Jan. 2019.

[26] A. J. Wimmers and C. S. Velden, "Objectively determining the rotational center of tropical cyclones in passive microwave satellite imagery," *J. Appl. Meteorol. Climatol.*, vol. 49, no. 9, pp. 2013–2034, Sep. 2010.

[27] A. J. Wimmers and C. S. Velden, "Advancements in objective multisatellite tropical cyclone center fixing," *J. Appl. Meteorol. Climatol.*, vol. 55, no. 1, pp. 197–212, Jan. 2016.

[28] N. Jaiswal and C. M. Kishtawal, "Automatic determination of center of tropical cyclone in satellite-generated IR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 3, pp. 460–463, May 2011.

[29] H. Akima, "A new method of interpolation and smooth curve fitting based on local procedures," *J. ACM (JACM)*, vol. 17, no. 4, pp. 589–602, Oct. 1970.

[30] C. S. Neumann and J. M. Pelissier, "An analysis of atlantic tropical cyclone forecast errors, 1970–1979," *Monthly Weather Rev.*, vol. 109, no. 6, pp. 1248–1266, Jun. 1981.

[31] B. R. Jarvinen, C. J. Neumann, and M. A. Davis, "A tropical cyclone data tape for the north atlantic basin, pp. 1886–1983: Contents, limitations, and uses," NOAA, Silver Spring, MD, USA, NOAA Tech. Memorandum NWS NHC 22, Mar. 1984. [Online]. Available: https://repository.library.noaa.gov/view/noaa/7069

[32] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization," *ACM Trans. Math. Softw.*, vol. 23, no. 4, pp. 550–560, Dec. 1997.

[33] J. Wertz and W. J. Larson, *Space Mission Analysis and Design*, 3rd ed. Amsterdam, The Netherlands: Springer, 1999.

Vinay Ravindra (Member, IEEE) received the M.Sc. degree in space technology from the Lulea University of Technology, Luleå, Sweden, the M.Sc. degree in space science and technology from the University of Würzburg, Würzburg, Germany, both in 2012, and the Ph.D. degree in electrical engineering and information systems from The University of Tokyo, Tokyo, Japan, in 2017.

He has worked as a Spacecraft Development Engineer at a Small-Satellite start-up based at Tokyo from 2012 to 2014 and from 2017 to 2018. From 2016 to 2017, he was a Project Research Associate with the Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency, Sagamihara, Japan. Since May 2018, he has been a Research Scientist within Earth Science Division, NASA Ames Research Center, Mountain View, CA, USA, contracted by Bay Area Environmental Research Institute, Petaluma, CA, USA, where his research focuses on optimizing Earth-science remote-sensing through the development of new observation strategies, instrument concepts, and related enabling technology.

Sreeja Nag (Member, IEEE) received the Ph.D. degree in space systems engineering from the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, USA, in 2015.

She is a Senior Research Scientist with the NASA Ames Research Center, Mountain View, CA, USA, contracted by the Bay Area Environmental Research Institute, Petaluma, CA, USA. Her research focuses on distributed vehicular robotics systems; specifically science-driven optimization of satellite constellations and autonomous operations, automated architectures for space traffic management, and safety-driven systems engineering for self-driving fleets.

Alan Li was born in Shanghai, China, in 1987. He received the B.S. degree in mechatronics engineering from the University of Waterloo, Waterloo, ON, Canada, in 2009, and the M.S. and Ph.D. degrees in aeronautics and astronautics engineering from Stanford University, Stanford, CA, USA, in 2011 and 2017, respectively.

In the Summer of 2011, he interned at Planet, San Francisco, CA, USA, with a focus on dynamic modeling of satellites. Since 2017, he has been a Research Engineer with NASA Ames Research Center, Mountain View, CA, USA, within the Laboratory for Advanced Sensing (LAS), Earth Sciences Division. His research interests include next-generation remote sensing technologies and machine learning as applied to remote sensing data sets.

Dr. Li is a member of the American Geophysical Union. He received the Outstanding Paper Award for Young Scientists at the 41st COSPAR Scientific Assembly in 2016.