1 **Assimilation of GOES-R Geostationary Lightning Mapper Flash Extent Density**

2 **data in GSI 3DVar, EnKF, and Hybrid En3DVar for the Analysis and Short-Term**

3 **Forecast of a Supercell Storm Case**

4

5 **Rong Kong[1], Ming Xue*[2], Edward R. Mansell[3], Chengsi Liu[1],**

6 **Alexandre O. Fierro[4]**

7 *[1] Center for Analysis and Prediction of Storms*

8 *[2] University of Oklahoma, Norman Oklahoma 73072*

9 *[3] NOAA/National Severe Storms Laboratory, Norman, Oklahoma*

10 *[4] Zentralanstalt für Meteorologie und Geodynamik, Department of Forecasting Models –*

11 *ZAMG. Vienna, Austria[1]*

12

13 **ABSTRACT**

14   Capabilities to assimilate Geostationary Operational Environmental Satellite "R-

15 series" (GOES-R) Geostationary Lightning Mapper (GLM) flash extent density (FED) data

16 within the operational Gridpoint Statistical Interpolation ensemble Kalman filter (GSI-

17 EnKF) framework were previously developed and tested with a mesoscale convective

18 system (MCS) case. This study further develops such capabilities to assimilate GOES GLM

19 FED data within the GSI ensemble-variational (EnVar) hybrid data assimilation (DA)

20 framework. The results of assimilating the GLM FED data using 3DVar, and pure

21 En3DVar (PEn3DVar, using 100% ensemble covariance and no static covariance) are

22 compared with those of EnKF/DfEnKF for a supercell storm case. The focus of this study

*Corresponding author: Ming Xue
Email: mxue@ou.edu

1

23   is to validate the correctness and evaluate the performance of the new implementation

24   rather than comparing the performance of FED DA among different DA schemes. Only

25   results of 3DVar and pEn3DVar are examined and compared with EnKF/DfEnKF.

26   Assimilation of a single FED observation shows that the magnitude and horizontal extent

27   of the analysis increments from PEn3DVar are generally larger than from EnKF, which is

28   mainly caused by using different localization strategies in EnFK/DfEnKF and PEn3DVar

29   as well as the integration limits of the graupel mass in the observation operator. Overall,

30   the forecast performance of PEn3DVar is comparable to EnKF/DfEnKF, suggesting

31   correct implementation.

32

33   **Key words:** GOES-R, lightning, convective-scale, data assimilation, EnKF, EnVar.

34   **Article Highlights:**

35   ● Implement GOES-R GLM FED data within the GSI ensemble-variational (EnVar)

36   hybrid data assimilation (DA) framework.

37   ● Forecast performance of PEn3DVar is comparable to EnKF but their analyses are

38   different, due to different localization strategies being used as well as the integration

39   limits of the graupel mass in the observation operator.

40

1. Introduction

High spatio-temporal resolution lightning data offer a reliable complement to radar data that can be assimilated into convection-allowing NWP models, especially in regions where the Next-Generation Radar (NEXRAD, Doviak et al., 2000) Weather Surveillance Radar-1998 Doppler (WSR-88D) radar coverage is poor or absent, such as in mountainous areas due to beam blockage, or in oceanic regions beyond the range of ground-based radars. Efforts have been made to assimilate lightning data into convection-allowing NWP models using nudging (e.g., Fierro et al., 2012; Marchand and Fuelberg, 2014; Fierro et al., 2015), 3DVar (Fierro et al., 2016; Fierro et al., 2019; Hu et al., 2020), 4DVar (Xiao et al., 2021) and EnKF methods (Mansell, 2014; Allen et al., 2016). Results show that assimilating either cloud-to-ground (CG), and especially total (CG plus intra-cloud) lightning data often produce forecast improvements comparable to those from assimilating ground-based radar data (Fierro et al., 2016). Among these studies, Mansell (2014) assimilated explicitly-simulated flash extent density (FED) data, and Allen et al. (2016) assimilated pseudo-geostationary lightning mapper (GLM, Goodman et al., 2013) observations derived from ground-based lightning mapping array (LMA, Rison et al., 1999; MacGorman et al., 2008) data, both using the more advanced ensemble Kalman filter (EnKF, Evensen, 1994) method. Here FED is a type of lightning flash rate and can be concisely defined as the total number of lightning flashes occurring within a given atmospheric column over a specified period (Lojou and Cummins, 2005). The readers are referred to Bruning et al. (2019) for more information on the derivation of GLM FED data. Observation operators for FED that depend on graupel echo volume or graupel mass were developed, based on output data from storm-scale simulations conducted with explicit electrification physics (Mansell et al., 2002,2005; Allen et al., 2016). With these observation operators, the FED is linearly

65   related to the total graupel mass or graupel volume within the FED pixel column. Other

66   studies by Fierro et al. (e.g., Fierro et al., 2012), Marchand and Fuelberg (2014) and Qie et

67   al. (2014) assimilated some form of pseudo data (water vapor mass, temperature, or ice-

68   phase particles) derived from the lightning data either through nudging or 3DVar, and such

69   assimilation was shown to improve convective initialization and short-term forecasts.

70   The GLM instruments on the GOES-R series geostationary satellites detect the light

71   emitted by lightning at the cloud tops and collects information about total lightning

72   discharges (Goodman et al., 2013). The GOES-R GLM is able to continuously detect total

73   lightning activity (intra-cloud plus CG) over the Americas and adjacent oceanic regions

74   with a spatial resolution/footprint ranging from 8-km near the center of the field of view

75   (near nadir) to about 12 km near the edges. GLM records information such as the

76   frequency, location, optical energy, and extent of light emitted from lightning discharges

77   to better depict the presence of developing and intensifying thunderstorms. NOAA/NASA

78   maintains two GOES-R series operational geostationary satellites: GOES-16 in the GOES-

79   East position and GOES-17 in the GOES-West position. Operational datasets from these

80   satellites have been made available since late 2017 and early 2019, respectively. In March

81   2022, GOES-18 was launched into space and went into operational service as GOES-West

82   on Jan. 4, 2023, replacing GOES-17.

83   Recently, real GLM lightning observations have been successfully assimilated (e.g.,

84   Fierro et al. (2019), Hu et al. (2020), and Kong et al. (2020)). Fierro et al. (2019) and Hu

85   et al. (2020) used indirect assimilation of lightning (flash density rates) through pseudo-

86   observations for water vapor mass derived from GLM lightning observations with a 3DVar

87   system (Gao et al., 2016; Wang et al., 2018). Kong et al. (2020) assimilated GLM lightning

88   (flash extent data) more directly using an EnKF method. In Fierro et al. (2019), the added

4

89   value of lightning DA was investigated by assimilating Level II radar data with and without

90   GLM-derived water vapor pseudo-observations for a group of severe convective storms.

91   The GLM and radar assimilation experiments yielded short-term forecast improvements

92   that were superior to those of a control run assimilating no data, with results on a par with

93   previous studies assimilating pseudo-GLM observations (Fierro et al., 2015; Fierro et al.,

94   2016).

95   Kong et al. (2020) was the first study to use an EnKF method to directly assimilate real

96   GOES-R lightning observations. In their proof-of-concept study, the capabilities to

97   assimilate GLM FED data were implemented into the NCEP operational EnKF DA system

98   (Gridpoint Statistical Interpolation, GSI, Kleist et al., 2009a) and tested for a mesoscale

99   convective system (MCS). FED observation operators based on graupel mass or graupel

100   volume were used, and the results were compared to a control experiment that did not

101   assimilate any data. The direct assimilation of FED observations was shown to noticeably

102   improve the analysis and prediction of the MCS for up to several hours. Akin to the indirect

103   methods, EnKF FED DA is effective in identifying and properly initializing regions of

104   intense convection, which helps improve shorter term forecasts of high impact convective

105   events. Direct and indirect adjustments to graupel and other model state variables through

106   the ensemble covariances helped produce more balanced analyses and, thus, improved

107   forecasts.

108   EnKF methods derive background error covariance from an ensemble of forecasts. The

109   hybrid ensemble-variational (EnVar) DA technique employs a combination of ensemble-

110   derived covariances and static covariances (typically used in 3DVar) to help alleviate the

111   problems with covariance matrix rank deficiency and other issues (Hamill and Snyder,

112   2000; Lorenc, 2003; Etherton and Bishop, 2004; Buehner et al., 2010). Hybrid EnVar

113   capabilities have been implemented within the NCEP operational GSI (Kleist et al., 2009b)

114   DA framework and are used operationally in the Global Data Assimilation System (GDAS,

115   Hamill et al., 2011; Kleist and Ide, 2015), the regional North American Mesoscale (NAM)

116   Data Assimilation System (Wu et al., 2017), and the Rapid Refresh (RAP, Benjamin et al.,

117   2016) DA system (Pan et al., 2014; Hu et al., 2017). For both the NAM and RAP, the

118   operational EnVar DA systems borrow ensemble perturbations from the GDAS EnKF

119   although coupling EnVar with one's own EnKF system is most desirable and has been

120   tested earlier for prototype RAP configurations (Pan et al., 2014; Pu et al., 2016; Lu et al.,

121   2017; Pan et al., 2018).

122   An earlier effort geared toward lightning data assimilation (LDA) using the hybrid

123   ensemble-variational DA method can be found in Apodaca et al. (2014). In their study,

124   pseudo-GLM data (derived from the ground-based lightning network) was assimilated

125   using a hybrid EnVar DA method and applied to mesoscale models in which deep

126   convection cannot be explicitly resolved. Their operator started with an approximate

127   calculation of vertical velocity from Nonhydrostatic Mesoscale Model core of the Weather

128   Research and Forecasting system (WRF-NMM), using a reduced version from the

129   nonhydrostatic continuity equation. Although the operator can analyze some updrafts that

130   are linked to other state variables through continuity equation, the relative coarser grid

131   resolution (e.g., 9 km in their study) is not suitable for convective scale DA.

132   Towards the eventual goal of operational implementation at NCEP, the convective-

133   scale EnKF DA capability for GLM FED data in Kong et al. (2020) was implemented

134   within the operational GSI EnKF framework. Given that hybrid EnVar is the current

135   method of choice for the two operational regional forecasting systems (NAM and RAP), it

136   is naturally desirable to implement and test hybrid EnVar DA capability for FED data

137   within the GSI framework. This study aims at validating the correctness and evaluating the

138   performance of the new implementation. In contrast to EnKF, EnVar obtains its analysis

139   by minimizing a cost function, as in 3DVar. In fact, it builds upon the GSI 3DVar

140   framework using the extended control variable algorithm of Lorenc (2003). The 3D version

141   of EnVar is referred to as En3DVar (Liu and Xue, 2016).

142   Specifically, this study includes three overarching goals: 1) to implement GLM FED

143   DA capabilities into the operational GSI EnVar that is coupled with GSI EnKF, 2) to

144   compare the performances of FED DA using EnKF and En3DVar to see whether

145   comparable results can be obtained between En3DVar and EnKF, and 3) to evaluate the

146   benefit of FED DA to short-range forecasts of a supercell storm case. Apart from lightning

147   and precipitation, we also evaluate the ability of FED DA to improve the updraft helicity

148   track forecast for the supercell storm case, which is commonly used as a proxy for tornado

149   prediction in operational centers (Kain et al., 2008; Sobash et al., 2016).

150   **2. Assimilation methods and procedures**

151   *2.1 The EnKF and DfEnKF algorithm*

152   The EnKF within GSI uses a scalable implementation of the ensemble square root filter

153   (EnSRF, Whitaker and Hamill, 2002), which updates state variables and pre-calculated

154   observation priors independently (Anderson and Collins, 2007). The scalable

155   implementation is equivalent to the traditional algorithm that updates state variables and

156   recalculate observation priors when using linear observation operators, and it can produce

157   qualitatively similar results as the traditional EnKF for practical NWP applications

158   (Anderson and Collins, 2007). For more detailed information about the updating equations

159   of the GSI EnKF algorithm, the reader is invited to consult Kong et al. (2020).

160   To facilitate more direct comparison between EnKF and pure En3DVar (PEn3DVar)

161   that uses 100% ensemble covariance (as in EnKF), a deterministic EnKF (DfEnKF)

162   algorithm introduced by Kong et al. (2018) is also adopted here. In DfEnKF, an additional

163   deterministic forecast (separate from the ensemble members) is updated using the same

164   EnKF equation that is used to update the ensemble mean. The updated state becomes the

165   initial condition for the deterministic forecast of the next cycle. The En3DVar is a

166   deterministic algorithm as well; it updates a single background forecast variationally,

167   utilizing the same ensemble perturbations as DfEnKF from a (one-way) coupled EnKF

168   system (see, e.g., Fig. 1 of Pan et al. 2014). Given that pure En3DVar is one-way coupled

169   with the EnKF, as is DfEnKF, the results between PEn3DVar and DfEnKF could be similar

170   but not necessarily the same. First, for EnKF/DfEnKF, observations are assimilated serially

171   (one after another), while for PEn3DVar, all observations are assimilated globally (at the

172   same time). Analyses could be different when the observation operator is nonlinear. Second,

173   for EnKF and DfEnKF, localizations are conducted between the model and the observation

174   space, while for En3DVar, localizations are conducted purely in the model space.

175   *2.2 The En3DVar algorithm*

176   The hybrid En3DVar algorithm based on the extended control variable method of

177   Lorenc (2003) is used in GSI, in which the full-rank static background error covariances

178   and the rank-deficient ensemble covariances are combined through the introduction of a

179   set of extended control vectors. The analysis increment is given by

180
$$\delta \mathbf{x_a} = \mathbf{x_a} - \mathbf{x_b} = \delta \mathbf{x_1} + \delta \mathbf{x_2}, \tag{1}$$

181   where $\delta \mathbf{x_1}$ and $\delta \mathbf{x_2}$ are the analysis increments related to the static and ensemble

182   background error covariances, respectively.

$$183 \qquad \delta \mathbf{x_2} = \sum_{i=1}^{N} \left(\mathbf{x}'_{bi} \circ \boldsymbol{\alpha}_i \right). \qquad (2)$$

184   Here $\boldsymbol{\alpha}_i$ is the extended control variable (vector); x'_{bi} is a vector of ensemble perturbations

185   (for the ith member) normalized by $\sqrt{N-1}$ with N being the ensemble size. $\mathbf{x}'_{bi} = \mathbf{x}_{bi} - \bar{\mathbf{x}}^{b}$

186   are the ensemble perturbations, and $\bar{\mathbf{x}}^{b} = \dfrac{1}{N} \sum_{i=1}^{N} \mathbf{x}_{bi}$ is the ensemble mean. The total analysis

187   increment δx_a is obtained by minimizing the following cost function:

$$188 \qquad \begin{aligned} J\left(\delta \mathbf{x}_1, \mathbf{a}\right) = {} & \frac{1}{2} \beta_1 \delta \mathbf{x}_1^{\mathrm{T}} \mathbf{B}^{-1} \delta \mathbf{x}_1 + \frac{1}{2} \beta_2 \boldsymbol{\alpha}^{\mathrm{T}} \mathbf{A}^{-1} \boldsymbol{\alpha} \\ & + \frac{1}{2} \left(\mathbf{H} \delta \mathbf{x}_1 + \mathbf{H} \sum_{i=1}^{N} \left(\mathbf{x}'_{bi} \circ \boldsymbol{\alpha}_i\right) + \mathbf{d} \right)^{\mathrm{T}} \mathbf{R}^{-1} \left(\mathbf{H} \delta \mathbf{x}_1 + \mathbf{H} \sum_{i=1}^{N} \left(\mathbf{x}'_{bi} \circ \boldsymbol{\alpha}_i\right) + \mathbf{d} \right) \end{aligned}, \qquad (3)$$

$$189 \qquad \frac{1}{\beta_1} + \frac{1}{\beta_2} = 1. \qquad (4)$$

190   where $1/\beta_1$ and $1/\beta_2$ are the weights given to the static and ensemble background error

191   covariances, respectively. A hybrid solution is derived by setting $1/\beta_1$ or $1/\beta_2$ to values

192   between 0 and 1. B is the static background error covariance matrix and A is the matrix

193   used to localize the ensemble background error covariance. d is the observation innovation

194   vector, H is the linearized observational operator, and R is the (diagonal) observation

195   error covariance matrix. When the weight for the static (ensemble) background error

196   covariance is 0, the hybrid algorithm degenerates to pure En3DVar (pure 3DVar).

197   In this study, temperature-dependent background error profiles for hydrometeor mixing

198   ratios proposed by Liu et al. (2019) are adopted when constructing the static background

199   error B to generate more reasonable graupel increments.

200   As in Kong et al. (2018), the hybrid En3DVar system is one-way coupled with EnKF.

201   With one-way coupling, it is easier to separate En3DVar from EnKF when comparing their

202   relative performance and understanding their respective behaviors. In general, little

203   difference is expected between one-way and two-way coupled EnKF-En3DVar when the

204   DA cycles are run for a short period (Pan et al., 2014). For long, continuously cycled DA

205   systems, however, two-way coupling is strongly recommended since divergence between

206   the EnKF and En3DVar systems can occur with time in a one-way coupling mode (Pan et

207   al., 2014).

208   *2.3 The FED observation operator*

209   The tuned FED observation operator used in this study is based on graupel mass instead

210   of graupel echo volume (Kong et al. (2020)). The graupel-volume-based FED observation

211   operator cannot be easily used in a variational framework, because of the zero gradient of

212   the cost function when differentiated with respect to the graupel mixing ratio. This is

213   because the graupel volume is no longer sensitive to graupel mixing ratio once its value

214   exceeds the threshold (e.g., 0.5 g kg^{-1}) used to define the graupel volume. The FED

215   observation operator based on graupel mass (GM in kg) from Kong et al. (2020) is given

216   by

$$\mathrm{FED} = 1.044 \times 10^{-8} \left(\mathrm{GM}\right). \tag{5}$$

217

218   This operator has been tuned through sensitivity experiments and the coefficient has been

219   multiplied by a factor of 0.5 compared to the original operator used in Allen et al. (2016).

220   Similar to Mansell (2014) and Allen et al. (2016), the graupel mass within each grid cell is

221   summed over a volume spanning the vertical depth of the model and covering a 15×15

222   km^2 horizontal area that is centered on the observation pixels, so that the lightning

223   occurring in the vicinity of a typical GLM 8×8 - 12×12 km^2 pixel can be accounted for.

10

224   Through the observation operator, information of lightning observations can be

225   transferred in the variational analysis to the graupel mixing ratio due to its direct linkage

226   with the FED observations. Other state variables (water vapor mixing ratio, temperature,

227   and vertical velocity, etc.) are linked through the ensemble-derived background error

228   covariances. When pure static background error covariance is used in 3DVar, only graupel

229   mixing ratio is directly updated by FED observations (due to lacking cross-variable

230   correlations in the static background error covariance B), although other state variables can

231   respond to the FED DA during the forecast period through microphysical or

232   thermodynamical adjustments.

233   The implementation of the FED observation operator involves the integration of the

234   graupel mass field over a fixed area (15×15 km^2 in our case) centered on the observation.

235   The number of model columns being used for integration is determined adaptively based

236   on model resolution. For example, a 5×5 region would be used for horizontal integration

237   on a 3-km model grid and 15×15 for the 1-km model grid used here.

238   **3. Simulation setup and experimental design**

239   *3.1 GOES-R GLM-FED data and their processing*

240   The FED observations were derived from the GOES-16 level-2 GLM data, which are

241   provided in 20-sec packets with an 8 - 12 km pixel resolution over the Contiguous United

242   States (CONUS, Goodman et al., 2013). The level-2 data include three different primary

243   lightning metrics - i.e., the flashes, groups and events - which are governed by specific

244   parent-to-child relationships (Goodman et al., 2013; Bruning et al., 2019). Similar to Kong

245   et al. (2020), the lightning observations are mapped onto 10-km grid pixels to obtain the

246   FED observations, with the number identification of the groups and flashes being tracked

11

247   to each matching event. The number of events is directly used to calculate the FED, which

248   by definition increments by one for each flash on each pixel (in contrast to flash source

249   densities). More details on the FED data processing can be found in Kong et al. (2020).

250   *3.2 Forecast model setup and initialization of ensembles*

251   The Advanced Research WRF (ARW) model is used in this study, which is a three-

252   dimensional compressible non-hydrostatic NWP model (Skamarock et al., 2008). The

253   model uses a 600×600 grid with 1-km horizontal grid spacing on a single domain and has

254   53 stretched vertical grid levels with a model top set at about 21 km.

255   Similar to the configurations of Kong et al. (2020), 40 initial ensemble members are

256   generated based on the deviations of the 2100 UTC analyses of the NCEP operational

257   Short-Range Ensemble Forecast system (SREF, which has 26 members with a 16-km

258   horizontal resolution) from the 3-h North American Model (NAM) forecast (with a 12-km

259   horizontal grid spacing) valid at the same time. The deviations are reduced by 25%

260   (determined through experimentation), and their negative versions are combined to

261   produce a total of 40 perturbations. Following Tong and Xue (2008), Snook et al. (2012),

262   and Johnson et al. (2014), additional smoothed small-scale random perturbations with

263   Gaussian distribution are added to the horizontal components of wind field, potential

264   temperature, and humidity fields to introduce convective-scale perturbations. The initial

265   ensembles are then advanced one hour to 2200 UTC, with the same surface and boundary

266   parameterizations and 6-class Thompson bulk microphysical scheme (Thompson et al.,

267   2008). More details on the physics options can be found in Kong et al. (2020).

268   *3.3 Design of data assimilation experiments*

269   As mentioned in section 2c, the graupel-mass-based observation operator from Kong

270   et al. (2020) was used in this study. The assimilation of GLM FED data using 3DVar,

271   EnKF, and pure En3DVar methods are evaluated against each other. Their results are also

272   compared with a control experiment (CTRL) that does not assimilate any data. Since the

273   purpose of this study is to validate the correctness of the implementation of FED DA within

274   the GSI variational DA framework, only 3DVar and pure En3DVar (instead of hybrid

275   En3DVar) are tested. More detailed information about different DA experiments is

276   provided in Table 1.

277   **Table 1**. Descriptions of the DA experiments.

Experiment	Use of background error covariance
CTRL (no DA)	N/A
EnKF	100% ensemble covariance
DfEnKF	100% ensemble covariance
3DVar	100% static & 0% ensemble covariance
PEn3DVar	0% static & 100% ensemble covariance

278

279   FED observations are assimilated in 5-min intervals over a 1-h period in all DA

280   experiments as in Kong et al. (2020). To help suppress spurious storms in the background,

281   zero FED observations are also assimilated. A 0.95 adaptive posterior inflation that relaxes

282   the posterior spread to the prior spread (Whitaker and Hamill, 2012; Kotsuki et al., 2017;

283   Maldonado et al., 2020) is used to help maintain ensemble spread. The FED operator is a

284   vertical integral of graupel mass, and, like the GLM FED observations, has no inherent

285   vertical location. Nevertheless, we need to specify the height of the FED observation for

286   vertical covariance localization in EnKF (which uses observation-space localization). The

287   FED observation is assigned to an assumed height of 6.5 km in the EnKF (Allen et al. 2016;

288   Kong et al. 2020), which generally falls within the mixed-phase region where most

13

289 lightning initiates. This is not needed in En3DVar because localization is conducted in the

290 model instead of observation space. Similar to Allen et al. (2016) and Kong et al. (2020),

291 the horizontal and vertical localization radii in the EnKF are set to 15 km and -4 in

292 $\log (P/P_{ref})$ space (~ 32 km on average). With the large vertical localization radius, there

293 is effectively no vertical localization, and the FED observations are allowed to influence

294 the whole vertical depth. The actual spatial spreading of analysis increments is governed

295 by the ensemble covariance (Kong et al., 2020). This is confirmed by single observation

296 experiment, where results are found to be insensitive to the specified height of the FED

297 observation (results not shown). In 3DVar and hybrid En3DVar, the decorrelation length

298 scale for the static background error covariances $\boldsymbol{B}$ is set to 6 km in the horizontal and to

299 1.5 grid points in the vertical (tuned through experiment), which are larger than those used

300 in Hu et al. (2020). Because the graupel mass observation operator is integrated over a 15

301 $\times 15\ km^2$ column, the FED observation does influence a much larger area than that implied

302 by the decorrelation length scale of $\boldsymbol{B}$. The horizontal and vertical decorrelation length

303 scales used to localize the ensemble covariances in En3DVar are 4.1 km and -1.1 in

304 $\log (P/P_{ref})$ space, respectively, which are equivalent to the cut-off radii of 15 km and -4

305 in $\log (P/P_{ref})$ used in EnKF based on Pan et al. (2014). Considering that the flash rates

306 from the GLM are overall smaller than those detected by very-high frequency ground-

307 based LMA (Rison et al., 1999) networks for high-flash-rate storms (Carey et al., 2019), a

308 small observation error of 0.5 flashes min^{-1} $pixel^{-1}$ is used in this study [as in Kong et al.

309 (2020)]. Moreover, initial sensitivity experiments revealed that setting an observation error

310 smaller than 0.5 flash min^{-1} $pixel^{-1}$ can result in observation overfitting and numerical

311 instability during model integration.

312   For the DA experiments, the first analysis is made at 2200 UTC of 1 May 2018, or 1 h

313   into the spinup ensemble forecasts from the perturbed initial conditions (Fig. 1). After the

314   1-h DA period, a 6-h forecast is made from the final analysis to examine the impact of
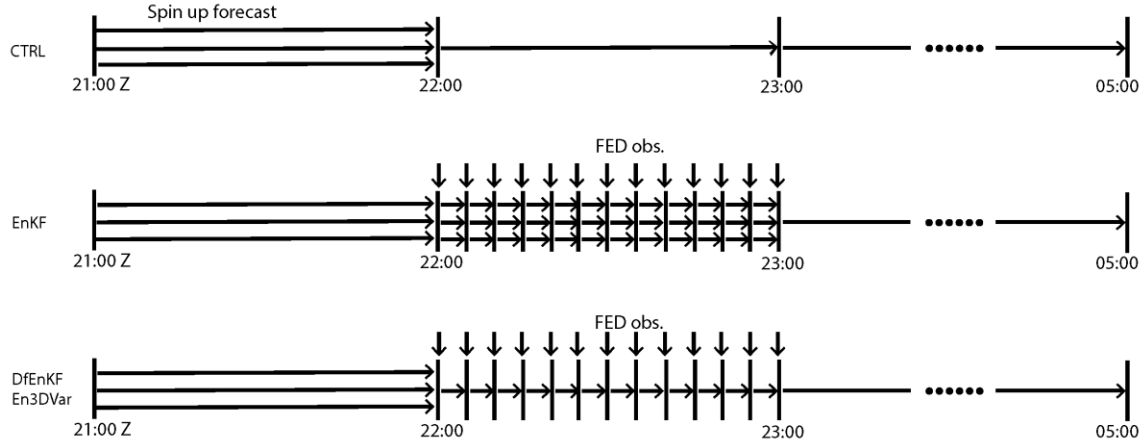
315   lightning DA on storm forecasts.

316

317

318   Fig. 1 Flow diagram of the EnKF (middle), 3DVar, and DfEnKF and En3DVar (bottom)
319   DA experiments versus the control run (top). The spin up ensemble forecasts from 2100
320   UTC through 2200 UTC include 40 members. FED DA occurs between 2200 and 2300
321   UTC with 5-minute cycles, and a deterministic forecast is launched from the final ensemble
322   mean analysis at 2300 UTC and run to 0500 UTC. The CTRL experiment continues the
323   ensemble forecasts through 2300 UTC without DA, and a single deterministic forecast
324   continues from the ensemble mean forecast at 2300 UTC.

325   *3.4 Case overview*

326   A severe weather event occurred in north central Kansas and southern Nebraska in the

327   afternoon and evening of May 1, 2018. Multiple supercells spawned nearly a dozen

328   tornadoes, including a long-track EF-3, across the state of Kansas. The storms were mainly

329   influenced by a stationary front over north-central Kansas (Fig. 2a-d), with favorable wind

330   shear and cold air aloft coupled with an abundant supply of low-level moisture advected

331   from the Gulf of Mexico by the low-level jet. A dryline extending southward from the

332   stationary front across west-central Kansas also triggered scattered thunderstorms. The

15

333 convective inhibition (CIN) of the 1200 UTC and 1900 UTC soundings from the NWS site

334 of Dodge City, Kansas (DDC) were 253 J Kg^{-1} and 4 J Kg^{-1}, respectively. The changes in

335 boundary-layer thermodynamics (reduction in inhibition vis erosion of the cap by solar

336 heating) coupled with a progressive destabilization of the environment (60 hPa mixed layer

337 CAPE at 1900 UTC exceeded 2000 J Kg^{-1}) allowed some cumulus turrets to develop along

338 the stationary front and along portions of the dry line in central Kansas. At 2100 UTC, a

339 line of semi-discrete storms developed along the stationary front (Fig. 2a). These cells

340 within the line progressively moved eastward while intensifying into severe-warned storms

341 by 2200 UTC (Fig. 2b). Our DA focuses on the period with frequent lightning activity

342 between 2200 UTC to 2300 UTC May 1st. Isolated supercell storms developed over the

343 next few hours, producing several tornadoes, strong wind events, and reports of hail up to
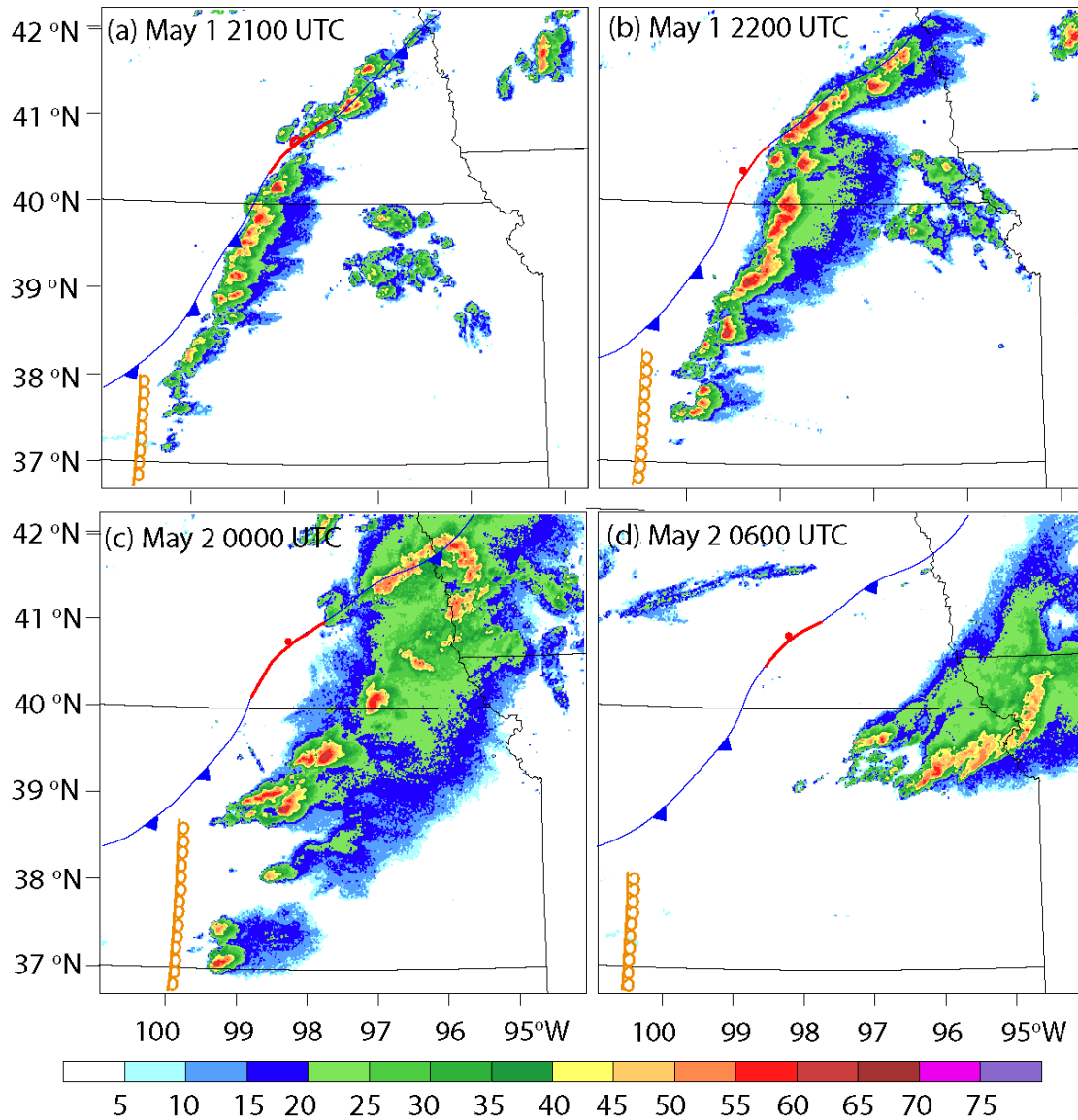
344 the size of softballs (10 cm diameter).

Fig. 2 Locations of the fronts (reproduced based on the surface analysis from the storm prediction center (SPC)), overlaid with the composite reflectivity observations (dBZ) remapped from the WSR-88D radars (shaded contour) on (a) 2100 May 1st, (b) 2200 May 1st, (c) 0000 May 2nd, and (d) 0600 UTC May 2nd, 2018, respectively.

4. Results

4.1 Convergence of the 3DVar/En3DVar cost function and computational costs of different DA methods.

354   The convergence of the quadratic cost function is evaluated first to validate the

355   correctness of the FED DA procedure in the variational DA framework. Since the

356   observation operator is linear, the double-loop procedure that is usually used to linearize a

357   nonlinear observation operator (such as for reflectivity) is not needed here. Only one outer

358   loop (with 50 maximum iterations) is used in this study. Figure 3a shows the value of the

359   cost function as a function of iterations in the first FED DA cycle for both cases. Pure

360   3DVar converges slightly faster than PEn3DVar. Overall, 15 iterations are sufficient for

361   convergence in both 3DVar and PEn3DVar. In our study, 50 iterations are used for all the

362   variational DA experiments to secure convergence of the cost function.

363   The computational costs of different DA schemes are also compared in Fig. 3b. Overall,

364   the computation cost of EnKF is 60% higher than 3DVar and both are much lower than

365   PEn3DVar, for which the cost function evaluation and minimization consume most of the

366   computational time. The EnKF computation/analysis is performed in the observational

367   space, and thus the computational cost is proportional to the number of observations. The

368   observational density of the two-dimensional FED data is much less than the three-

369   dimensional radar data. If FED data are assimilated together with 3D radar data in EnKF,

370   the additional cost incurred by the FED DA becomes negligible. For EnVar, the

371   computational cost is proportional to the length of state vector, which includes the much

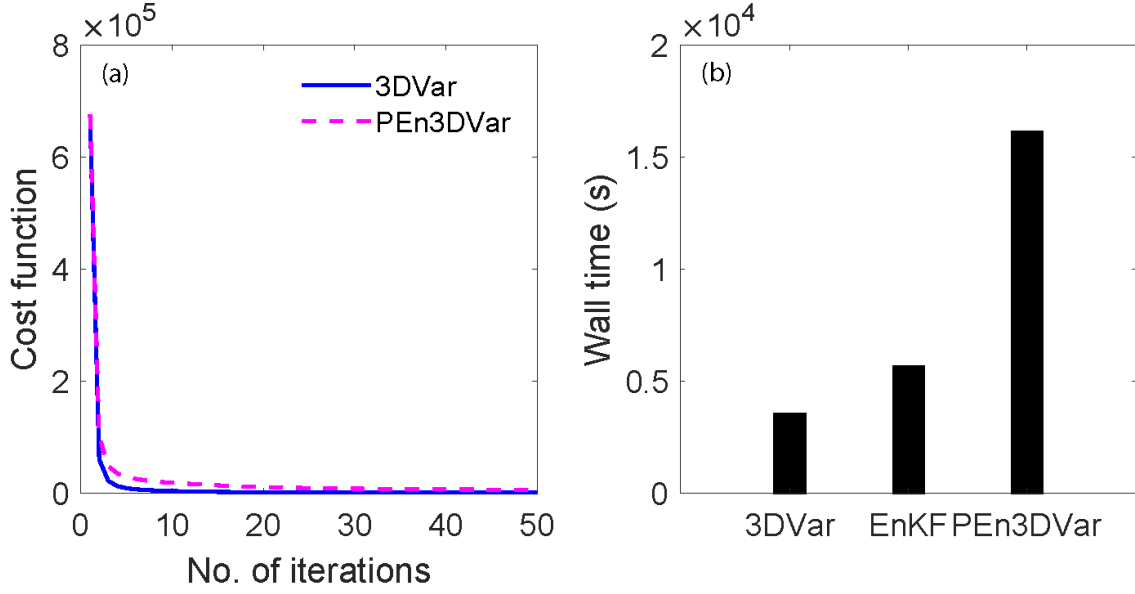372   larger extended state vector α (see Eq. (2)).

373

Fig. 3 (a) Value of the cost function as a function of iteration loops for both the 3DVar and PEn3DVar runs (there is only one outer loop, with 50 iterations in the loop) in the 1st analyses and (b) the computational cost (wall time in seconds) of 3DVar, EnKF, and PEn3DVar.

4.2 Single point observation experiment

To demonstrate the direct impacts of lightning DA on the analyses of hydrometeor fields, single point experiments (with grid column x=230, y=360 being the closet to the center of the single observation pixel) are conducted and compared between DfEnKF and pure En3DVar that uses 100% ensemble covariance (PEn3DVar). Vertical cross-sections of the analysis increments [$\mathbf{x_a} - \mathbf{x_b}$ in Eq. (1)] for mixing ratios of rain, snow, and graupel are made through the single observation point. As in Kong et al. (2020), the horizonal localization radius for the DfEnKF is 15 km (15-km is the zero point) and the vertical localization scale is 4 in logP space. Since the vertical localization radius is very large (about to 32 km on average in vertical), effectively the vertical location is not considered in our DA experiments. To ensure similar localization scales being used in both PEn3DVar and EnKF, equivalent recursive filter length scales based on Pan et al. (2013) are used for localization, which is about 4.11 km in horizontal and 1.10 in logP space for PEn3DVar.

19

391   As is shown in Fig. 4, the magnitude and horizontal extent of the analysis increment of

392   the hydrometeors of graupel, snow, and rainwater for PE3DVar (Fig. 4b) are much larger

393   than DfEnKF (Fig. 4a). The differences in the analysis increments probably arise from the

394   localization and the volume integration of the graupel mass in the observation operator. To

395   eliminate the influence caused by horizontal integration of the graupel mass in the FED

396   observation operator, the horizontal integration is arbitrarily turned off and the vertical

397   integration remains unchanged. In other words, the integration is conducted only over a

398   single grid column with a 1 by 1 km^2 area and multiplying by a factor of 15^2 (i.e., the full

399   area) to achieve similar physical meaning as the original operator. The observation operator

400   is only modified to help understand the problem, and the original version was used for the

401   experiments discussed in the following sections. After turning off the horizontal integration

402   in the observation operator, the horizontal extension of the analysis increment from

403   DfEnKF (Fig. 4c) and PEn3DVar (Fig. 4d) becomes very similar. Thus, the horizontal

404   integration in the observation operator is the major cause for the difference in the analysis

405   increment between DfEnKF and En3DVar under the influence of different localization
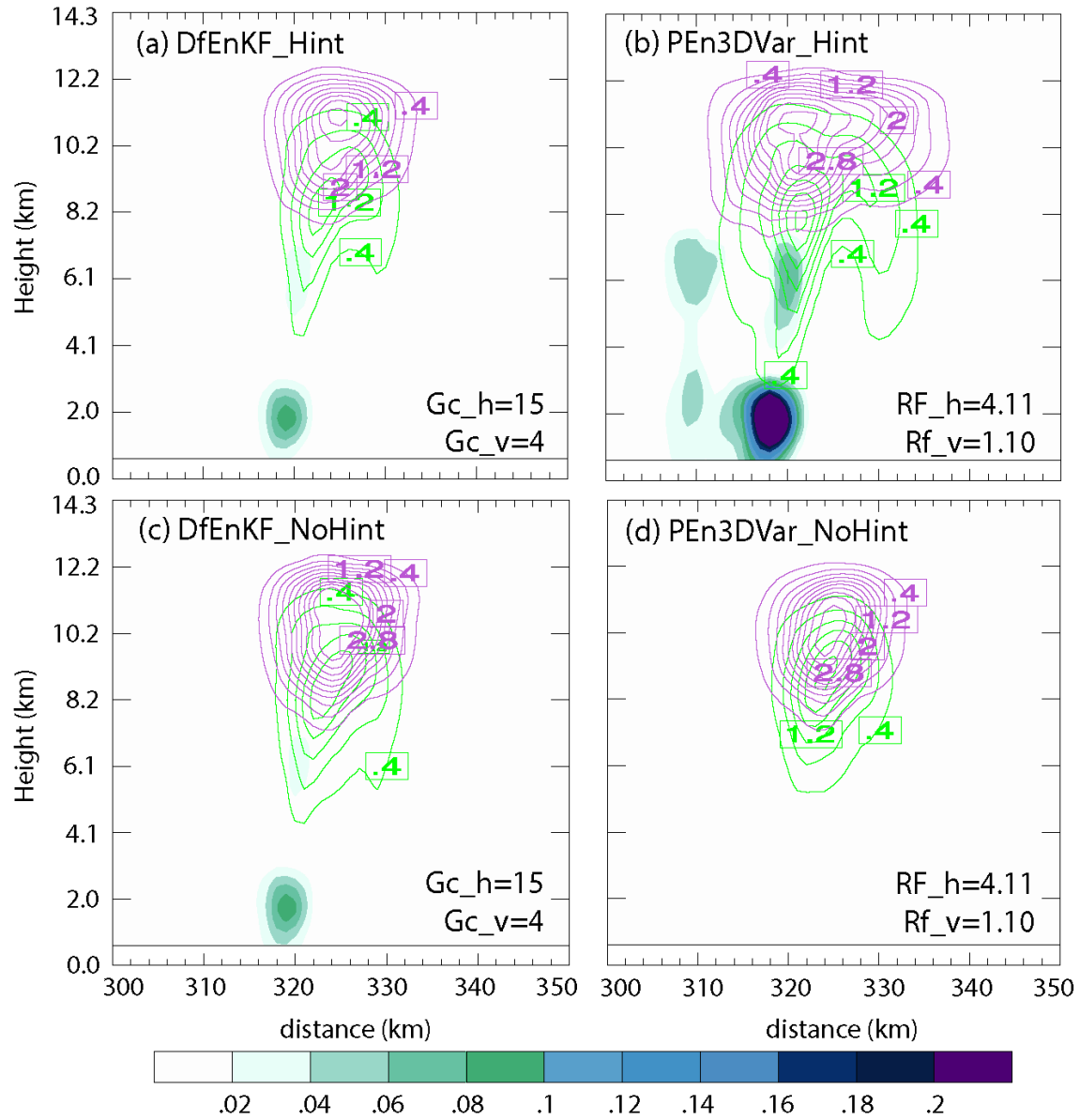
406   strategies.

Fig. 4 Vertical cross sections of the mixing ratio increments (g kg^{-1}) of rain (shading), snow (purple contours, start from 0.4 with 0.4 g kg^{-1} interval), and graupel (green contours, start from 0.4 with 0.4 g kg^{-1} interval) after the first analysis of (a, c) DfEnKF and (b, d) PEn3DVar using (a, b) the original observation operator and (c, d) the modified observation operator, respectively. The cross section passes through the observation point (x=230 km, y=360 km).

4.3 Sensitivity of DfEnKF and En3DVar analyses to different localization scales

To further understand the impacts of the difference in localization and integration volume in the observation operator, sensitivity experiments are conducted to assimilate

21

417   single FED observation with different horizontal and vertical localization scales. Unlike

418   En3DVar, for which localization is conducted in the model space, EnKF needs the height

419   information of the observation to conduct localization between model and observation

420   space. Following Allen et al. (2016) and Kong et al. (2020), the FED observations are

421   assigned to a height of 6.5 km for DfEnKF DA. When vertical localization radius is small

422   (0.2 in log(P) space), the difference between DfEnKF and PEn3DVar is most obvious. As

423   is shown in Fig. 5a, with small vertical localization radius, the analysis increments of

424   graupel field for DfEnKF only exists in the levels round 6.5 km, while the analysis

425   increment of PEn3DVar has more extension in vertical, since localization of En3DVar is

426   conducted in the model space. When increasing the vertical location radius but keeping the

427   horizontal localization the same, the difference between En3DVar and DfEnKF becomes

428   smaller (Figs. 5b, c, f, g). When also increasing the horizontal localization radius to a very

429   large value (e.g., 2000 km cutoff for DfEnKF), the difference between DfEnKF and

430   PEn3DVar becomes very small. Without localization in both the horizontal and vertical

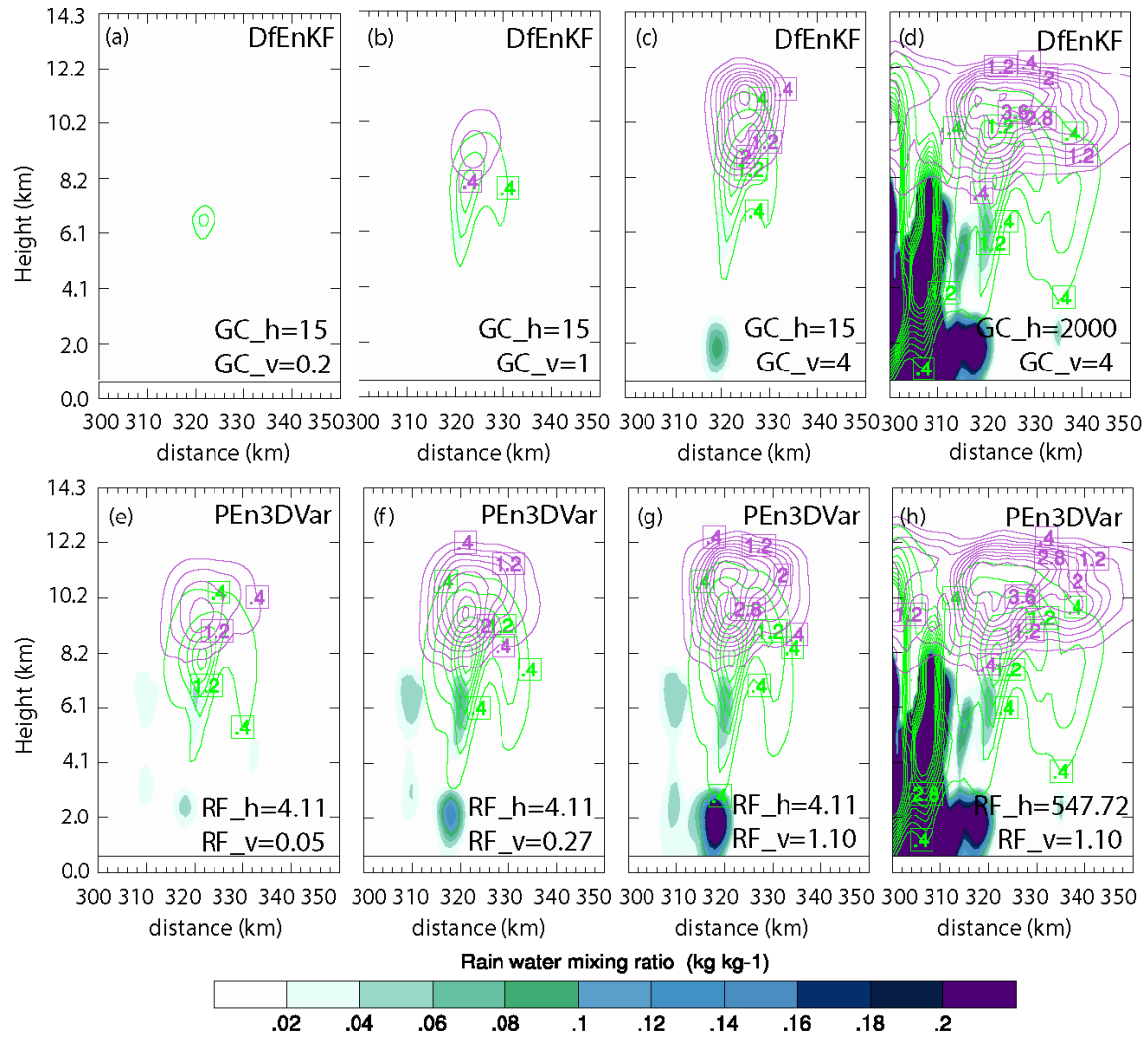431   directions, DfEnKF and PEn3DVar can get similar analysis increments (Fig. 5d, h).

Fig. 5 Vertical cross sections of the mixing ratio increment (g kg^{-1}) of rain (shading), snow (purple contours, start from 0.4 with 0.4 g kg^{-1} interval), and graupel (green contours, start from 0.4 with 0.4 g kg^{-1} interval) after the first analysis of (a, b, c, d) DfEnKF with horizontal localization of (a, b, c) 15 km (d) 2000 km and vertical localization radius of (a, b) 0.2, 1, (c, d) 4 in log(P) space, and (e, f, g, h) PEn3DVar with horizontal recursive filter length scale of (e, f, g) 4.11 km and (h) 547.72 km and vertical localization of (a) 0.05, (b) 0.27, and (c, d) 1.10 in log(P) space, respectively. The cross section passes through the observation point (x=230 km, y=360 km).

4.4 Comparisons of EnKF, DfEnKF, 3DVar, and PEn3DVar

To evaluate how well the analyses match the observations after 1-h DA, the FED observation operator is applied to the final analysis of the DA experiments and the simulated FEDs are compared with the FED observations. For the FED analyses after 1-h DA, EnKF, DfEnKF, and PEn3DVar are all able to better capture the observed discrete

23

446 FED cores corresponding to the three main storm cells, which are more consistent with the

447 observations than 3DVar or CTRL (Fig. 6). The distribution of the FED field from CTRL

448 is too linear and there are obvious location errors. 3DVar produces a poorer FED analysis

449 with more small-magnitude FEDs relative to EnKF, DfEnKF, and PEn3DVar as well as

450 the observations, which is expected since 3DVar only updates graupel field instead of the

451 whole model states as in EnKF/DfEnKF and PEn3DVar during DA.

452

453
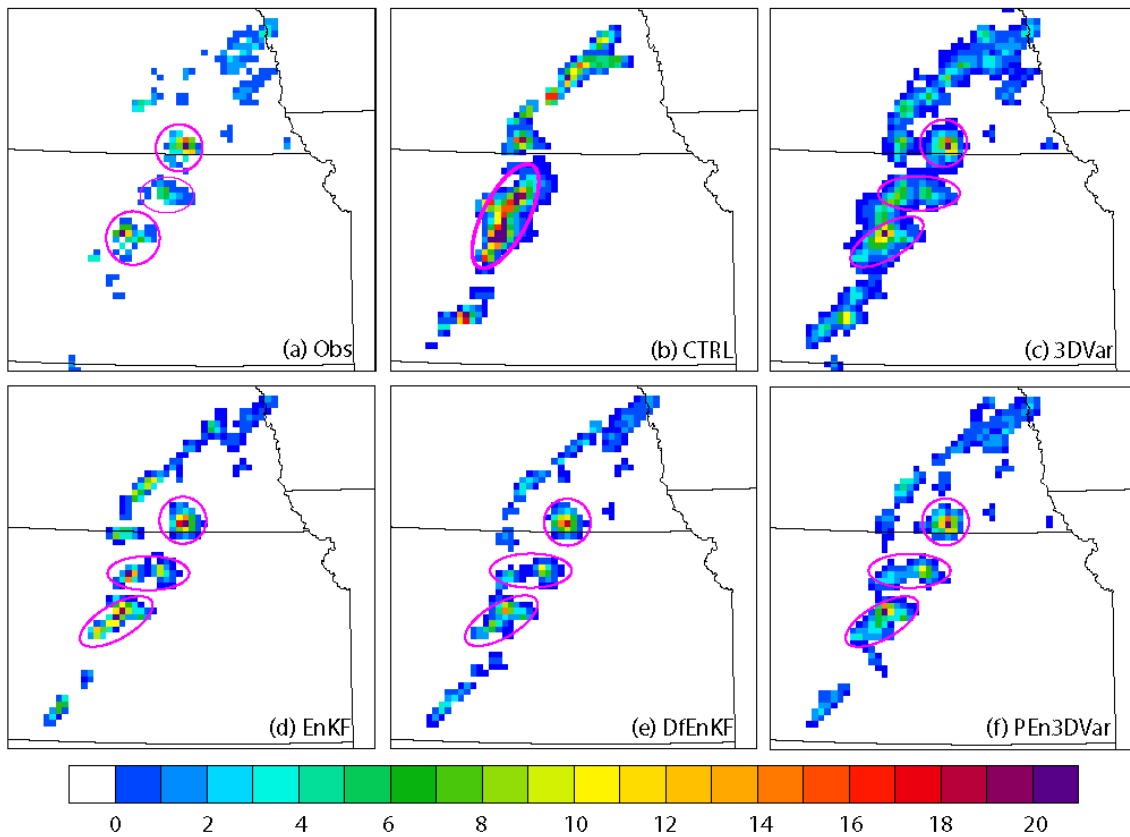
454 Fig. 6 (a) 1-min FED observations (units: flash min^{-1} $pixel^{-1}$) and FED forecasts from (b)
455 CTRL; analyses (after 1-h DA) from (c) 3DVar; (d) EnKF, (e) DfEnKF, and (f) PEn3DVar.

456   To quantitatively assess the performance of different DA experiments, the root-mean-

457 square innovation (RMSI)s of FED analyses and forecasts within the 1-h DA window and

458 the follow-on 1-4 h FED forecasts for CTRL, 3DVar, EnKF, DfEnKF, and PEn3DVar are

459   displayed in Fig. 7. Overall, EnKF performs best and has fewer spurious FED areas relative

460   to the other DA methods (Figure not shown). Within the DA cycles, taking the average of

461   ensemble forecasts to obtain ensemble means can lead to weaker FED values and thus has

462   less chance of overestimation (producing less spurious storms, figures not shown) in the

463   EnKF analyses and forecasts relative to DfEnKF and PEn3DVar (which does not involve

464   ensemble mean). 3DVar has smaller RMSI than CTRL during DA, but increases quickly

465   during the 1-h forecasts after the DA. The lack of consistent updating of all state variables
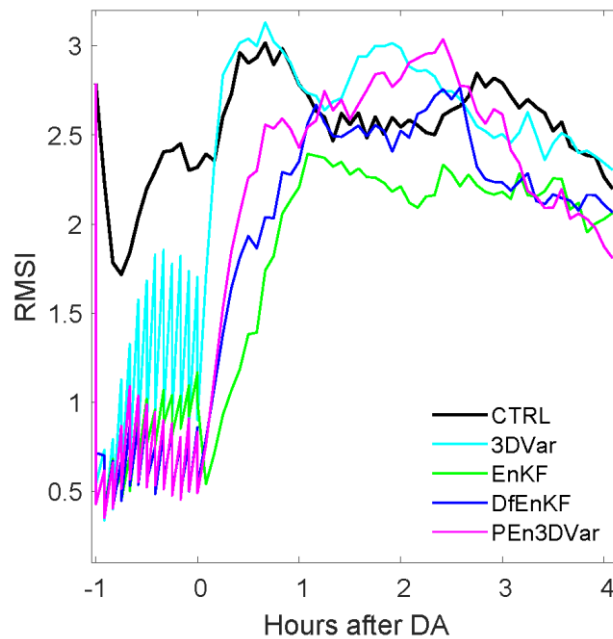
466   is believed to be the main reason.

467

468

Fig. 7 RMSIs of the FED forecasts (flash min^{-1} pixel^{-1}) for CTRL; FED analyses and
forecasts within the 1-h DA window, and 0-4 h free forecasts for 3DVar, EnKF, DfEnKF,
and PEn3DVar, respectively.

472   To quantitively evaluate how well different DA experiments capture the structure and

473   intensity of the storms, the 0 - 4 h forecasts of composite reflectivity fields from different

474   experiments are verified against WSR-88D composite reflectivity using the equitable threat

475   score (ETS; Mason, 2003) with 5-km neighborhood radius in Fig. 8. For the threshold of

476   20 dBZ (Fig. 8a), EnKF outperforms DfEnKF and PEn3DVar, apparently because EnKF

477   generates fewer spurious storms than DfEnKF and PEn3DVar (figures not shown). Both

478   DA methods produce overall better forecasts than 3DVar, which in turn still outperforms

479   CTRL (in terms of ETS). For the threshold of 35 dBZ (Fig. 8b), EnKF produces the best

480   0-1-h forecast. The ETSs of DfEnKF, PEn3DVar and 3DVar are similar and drop quickly

481   during the first 0.5 h of forecast. For 1 to 3 h, the DA experiments do not exhibit substantial
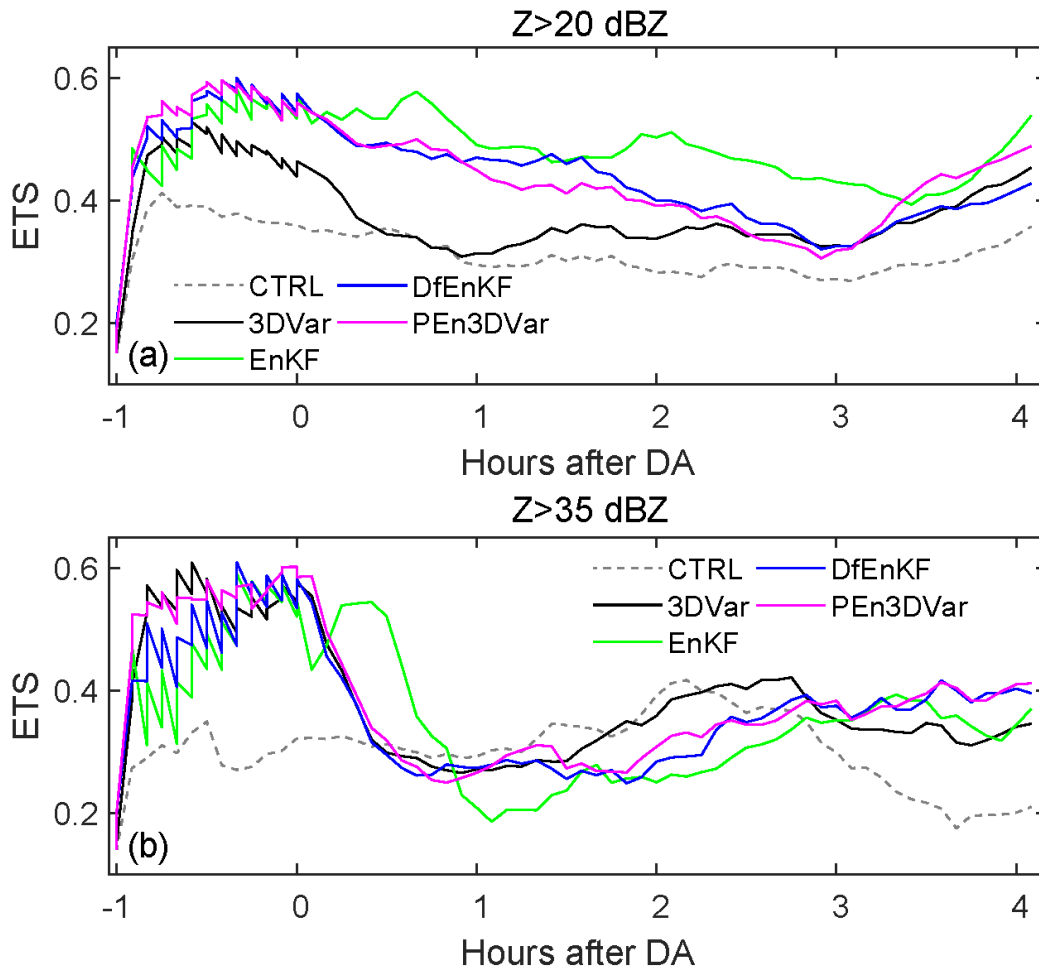
482   improvements over CTRL.

483

485   Fig. 8 5-km neighborhood ETSs of the supercell composite reflectivity analyses and
486   forecasts verified against (a) 20 dBZ and (b) 35 dBZ within 1-h DA (corresponding to -1
487   to 0 h after DA) and 0 - 4 h composite reflectivity forecasts after DA for experiments
488   CTRL, 3DVar, EnKF, DfEnKF, and PE3DVar, respectively.

489   Updraft helicity (UH; Kain et al., 2008) is an integrated measure of updraft rotation in

490   supercells, and UH swaths have been shown to be a good proxy for tornado track

491   forecasting (Clark et al., 2012). To evaluate how well different DA experiments predict

492   rotating updrafts and tornado potential, the accumulated swaths of the $2 - 5$ km UH from

493   the 0-2 forecasts of the velocity fields are compared with the locations of official Storm

494   Prediction Center (SPC) tornado and hail reports. As can be seen in Fig. 9, EnKF, DfEnKF,

495   and pure En3DVar perform overall similarly and outperform both CTRL and 3DVar in

496   terms of capturing the three different high UH swaths in southern Nebraska and north

497   central Kansas that are consistent with the tracks of tornado and hail reports. The ensemble-

498   based DA methods (EnKF/DfEnKF or En3DVar) all help improve the forecast location of

499   rotating updrafts, and thus are able to provide more accurate prediction of potential severe

500   weather threats for this case.

501
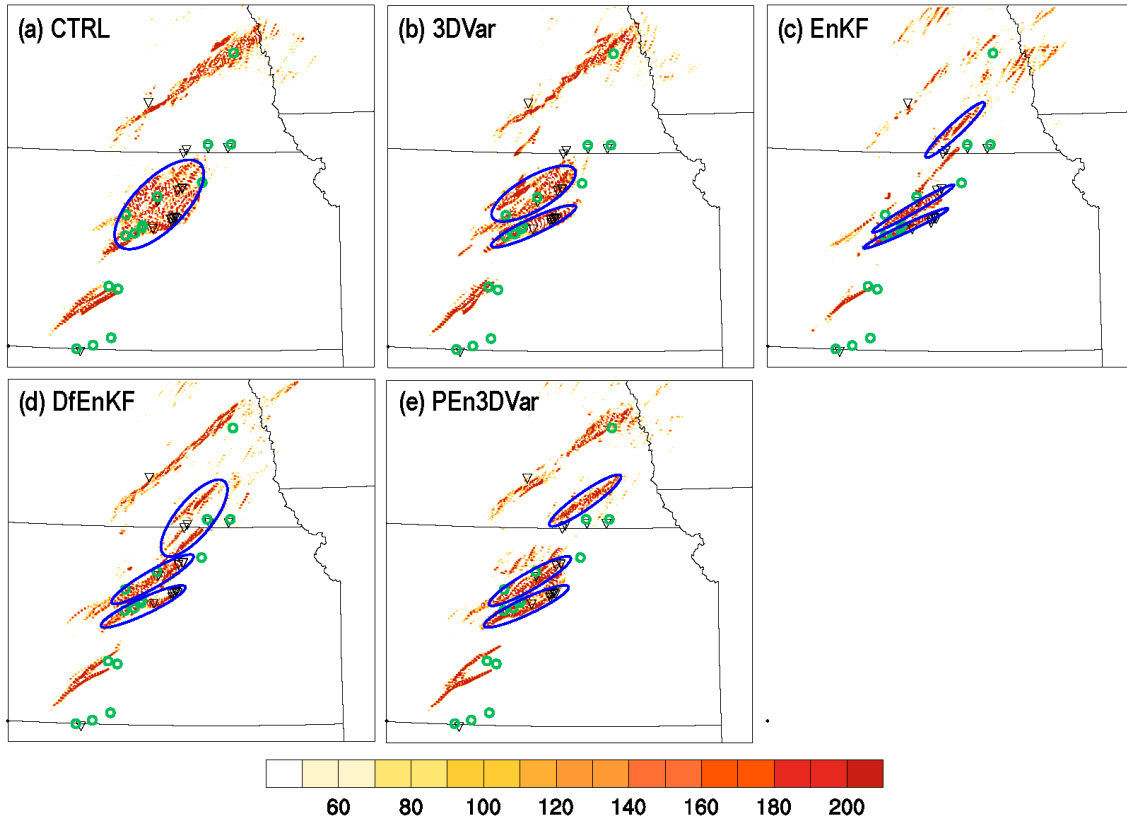
Fig. 9 The 0 - 2 h forecasted 2-5 km updraft helicity ($m^2 s^{-2}$) for the supercell case from (a) CTRL, (b) 3DVar, (c) EnKF, (d) DfEnKf, and (e) PEn3DVar, respectively, overlaid with tornado reports (inverted triangle) and hail reports (green circles) during the same time period.

507

5. Summary and Conclusions

In this study, capability was developed to assimilate GOES-R GLM data within the GSI hybrid ensemble-variational (EnVar) system for convection-allowing NWP. Real GOES-R GLM FED data are assimilated into convection-resolving WRF model for a supercell storm case over the Great Plains of United States using the hybrid En3DVar method, EnKF, as well as the DfEnKF and, 3DVar methods. The FED observation operator based on graupel mass from Kong et al. (2020) is used for the assimilation. 1-min FED

28

515　rates are assimilated every 5 min for a 1-h period and 4-h deterministic forecasts are

516　produced from the final analyses of each DA method.

517　　　DA experiments using EnKF, DfEnKF, 3DVar, and hybrid En3DVar are performed,

518　and their performances are evaluated and compared against each other as well as against a

519　control run (CTRL) that does not assimilate any data. Objective verifications employing

520　neighborhood ETS are performed for simulated FED and reflectivity. The main findings

521　from this evaluation are summarized as follows:

522　　• Single point experiments are conducted and compared between DfEnKF and pure

523　　　En3DVar at uses 100% ensemble covariance (PEn3DVar). Without applying

524　　　vertical localization in both En3DVar and EnKF, the magnitudes and horizontal

525　　　extents of the analysis increments of the hydrometeors of graupel, snow, and

526　　　rainwater for PE3DVar are much larger than EnKF. The differences in the analysis

527　　　increments are mainly caused by the localization and the horizontal/vertical

528　　　integration of the graupel mass in the observation operator.

529　　• The analyses after 1-h DA and the ensuing 4-h forecasts of FED in EnKF, DfEnKF,

530　　　3DVar, and PEn3DVar are compared with those of CTRL as well as GLM FED

531　　　observations. All DA experiments generally capture the three discrete FED cores

532　　　better than CTRL in the analyses. The high FED values in CTRL appear in more

533　　　linear structure and lack discrete FED core structures as observed.

534　　• The root-mean-square innovation (RMSI)s of FED analyses and forecasts within

535　　　the 1-h DA window and the follow-on 1-4 h FED forecasts for CTRL, 3DVar,

536　　　EnKF, DfEnKF, and PEn3DVar are compared. Overall, pure En3DVar that uses

537　　　100% ensemble covariance has similar performance as EnKF/DfEnKF. 3DVar has

538　　　smaller RMSI than CTRL during DA, but increases quickly during the 1-h forecasts

539   after the DA. The lack of consistent updating of all state variables is believed to be

540   the main reason.

541   • The 0 - 4 h forecasts of composite reflectivity fields from different experiments are

542   verified against WSR-88D composite reflectivity. Overall, all DA experiments

543   produce higher ETSs than CTRL; EnKF produces slightly higher ETSs than

544   DfEnKF and PEn3DVar and both are much larger than for 3DVar for the threshold

545   of 20 dBZ.

546   • The 0-2 h accumulated swaths of 2 – 5 km UH in the supercell DA experiments are

547   compared with SPC tornado reports. EnKF, DfEnKF, and PEn3DVar all perform

548   similarly and outperform both CTRL and 3DVar in terms of capturing the three

549   high UH swaths in central Kansas matching the observed tornado tracks.

550   In summary, the experiments that assimilate FED data using EnKF, 3DVar, and

551 En3DVar are all better at capturing the intensity and distribution of storms in the analyses

552 and forecasts in terms of FED, reflectivity, and updraft helicity relative to CTRL that does

553 not assimilate FED data. Overall, PEn3DVar produces comparable results as EnKF,

554 whereas 3DVar performs the worst among all DA methods tested. This suggests that the

555 flow-dependent cross-variable correlations used in ensemble-based DA experiments

556 (EnKF, En3DVar) play an important role in generating physically consistent analyses of

557 unobserved state variables, leading to improved forecasts.

558   Despite encouraging results obtained thus far within the operational GSI DA

559 framework, more systematic testing and evaluations are warranted with more cases. The

560 current operators assumed a linear relationship between FED and graupel mass or graupel

561 volume and the operators were originally derived based on simulation data sets. The

562 relationship between total flash rate and graupel mass is not necessarily linear (e.g., Fig.

563 5a in Allen et al., 2016), and there may be large variability in the graupel mass scaling

564 coefficient across different convective regimes. Kong et al. (2022) recently developed a

565 new nonlinear operator for the GLM FED observations, which can be tested and compared

566 with the old operator using variational DA methods in future work. Furthermore, the GLM

567 FED data would eventually need to be assimilated in conjunction with other operationally

568 available data, such as radar and GOES-R ABI radiance data. The sensitivity of data impact

569 to microphysics parameterization scheme should also be investigated as it directly affects

570 thunderstorm prediction. In addition, the new DA methods can also be implemented in

571 Joint Effort for DA Integration (JEDI, which is an innovative, next-generation U.S. DA

572 system) and tested with the target Rapid Refresh Forecast System (RRFS, the next-

573 generation U.S. ensemble weather forecast system) configurations. It is hoped that after

574 more intensive testing, the capabilities can be transitioned into operations in the near future.

575

587

588 **REFERENCES**

589 Allen, B. J., E. R. Mansell, D. C. Dowell, and W. Deierling, 2016: Assimilation of Pseudo-
590   GLM Data Using the Ensemble Kalman Filter. *Mon. Wea. Rev.*, **144**, 3465-3486.
591 Anderson, J. L., and N. Collins, 2007: Scalable implementations of ensemble filter
592   algorithms for data assimilation. *J. Atm. Ocean. Tech.*, **24**, 1452-1463.
593 Apodaca, K., M. Zupanski, M. DeMaria, J. A. Knaff, and L. D. Grasso, 2014: Development
594   of a hybrid variational-ensemble data assimilation technique for observed lightning
595   tested in a mesoscale model. *Nonlinear Proc Geoph*, **21**, 1027-1041.
596 Benjamin, S. G., and Coauthors, 2016: A North American Hourly Assimilation and Model
597   Forecast Cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669-1694.
598 Bruning, E. C., and Coauthors, 2019: Meteorological Imagery for the Geostationary
599   Lightning Mapper. *J. Geophys. Res. Atmos.*, **124**, 14285-14309.
600 Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010:
601   Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter
602   for Global Deterministic NWP. Part II: One-Month Experiments with Real
603   Observations. *Mon. Wea. Rev.*, **138**, 1567-1586.
604 Carey, L. D., N. Curtis, S. M. Stough, C. J. Schultz, and P. M., 2019: A Radar Investigation
605   of Storm Morphology and Precipitation Processes during Discrepancies between
606   GOES-16 GLM and LMA Observed Lightning Flash Rates and Jumps. 15th Annual
607   Symposium on New Generation Operational Environmental Satellite Systems, 6-
608   10 Jan. *Bull. Amer. Meteor. Soc.*, Phoenix, AZ.
609 Clark, A. J., J. S.Kain, P. T.Marsh, J.CorreiaJr., M.Xue, and F.Kong, 2012: Forecasting
610   tornado pathlengths using a three-dimensional object identification algorithm
611   applied to convection-allowing forecasts. . *Wea. Forecasting*, **27**, 1090–1113.
612 Doviak, R. J., V. Bringi, A. Ryzhkov, A. Zahrai, and D. Zrnic, 2000: Considerations for
613   polarimetric upgrades to operational WSR-88D radars. *J. Atm. Ocean. Tech.*, **17**,
614   257-278.
615 Etherton, B. J., and C. H. Bishop, 2004: Resilience of hybrid ensemble/3DVAR analysis
616   schemes to model error and ensemble covariance error. *Mon. Wea. Rev.*, **132**, 1065-
617   1080.
618 Evensen, G., 1994: Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic
619   Model Using Monte-Carlo Methods to Forecast Error Statistics. *J. Geophys. Res-
620   Oceans.*, **99**, 10143-10162.
621 Fierro, A. O., E. R. Mansell, C. L. Ziegler, and D. R. MacGorman, 2012: Application of a
622   Lightning Data Assimilation Technique in the WRF-ARW Model at Cloud-
623   Resolving Scales for the Tornado Outbreak of 24 May 2011. *Mon. Wea. Rev.*, **140**,
624   2609-2627.
625 Fierro, A. O., Y. Wang, J. Gao, and E. R. Mansell, 2019: Variational Assimilation of Radar
626   Data and GLM Lightning-Derived Water Vapor for the Short-Term Forecasts of
627   High-Impact Convective Events. *Mon. Wea. Rev.*, **147**, 4045-4069.
628 Fierro, A. O., A. J. Clark, E. R. Mansell, D. R. MacGorman, S. R. Dembek, and C. L.
629   Ziegler, 2015: Impact of Storm-Scale Lightning Data Assimilation on WRF-ARW
630   Precipitation Forecasts during the 2013 Warm Season over the Contiguous United
631   States. *Mon. Wea. Rev.*, **143**, 757-777.
632 Fierro, A. O., J. Gao, C. L. Ziegler, K. M. Calhoun, E. R. Mansell, and D. R. MacGorman,
633   2016: Assimilation of Flash Extent Data in the Variational Framework at

Convection-Allowing Scales: Proof-of-Concept and Evaluation for the Short-Term Forecast of the 24 May 2011 Tornado Outbreak. *Mon. Wea. Rev.*, **144**, 4373-4393.

Gao, J., C. Fu, J. Kain, and D. J. Stensurd, 2016: OSSE experiments for an ensemble of 3DVAR data assimilation system with radar data for convective storms. *J. Atnos. Sci.*, **73**, 2403-2426.

Goodman, S. J., and Coauthors, 2013: The GOES-R Geostationary Lightning Mapper (GLM). *Atmos. Res.*, **125**, 34-49.

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter-3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905-2919.

Hamill, T. M., J. S. Whitaker, D. T. Kleist, M. Fiorino, and S. G. Benjamin, 2011: Predictions of 2010's Tropical Cyclones Using the GFS and Ensemble-Based Data Assimilation Methods. *Mon. Wea. Rev.*, **139**, 3243-3247.

Hu, J., A. O. Fierro, Y. Wang, J. Gao, and E. R. Mansell, 2020: Exploring the assimilation of GLM derived water vapor mass in a cycled 3DVAR framework for the short-term forecasts of high impact convective events. *Mon. Wea. Rev.*, **148**, 1005-1028.

Hu, M., S. G. Benjamin, T. T. Ladwig, D. C. Dowell, S. S. Weygandt, C. R. Alexander, and J. S. Whitaker, 2017: GSI Three-Dimensional Ensemble-Variational Hybrid Data Assimilation Using a Global Ensemble for the Regional Rapid Refresh Model. *Mon. Wea. Rev.*, **145**, 4205-4225.

Johnson, A., and Coauthors, 2014: Multiscale Characteristics and Evolution of Perturbations for Warm Season Convection-Allowing Precipitation Forecasts: Dependence on Background Flow and Method of Perturbation. *Mon. Wea. Rev.*, **142**, 1053-1073.

Kain, J. S., and Coauthors, 2008: Some Practical Considerations Regarding Horizontal Resolution in the First Generation of Operational Convection-Allowing NWP. *Wea. Forecasting*, **23**, 931-952.

Kleist, D. T., and K. Ide, 2015: An OSSE-Based Evaluation of Hybrid Variational-Ensemble Data Assimilation for the NCEP GFS. Part II: 4DEnVar and Hybrid Variants. *Mon. Wea. Rev.*, **143**, 452-470.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W. Wu, and S. Lord, 2009a: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea. Forecasting*, **24**, 1691-1705.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. H. Yang, 2009b: Improving Incremental Balance in the GSI 3DVAR Analysis System. *Mon. Wea. Rev.*, **137**, 1046-1060.

Kong, R., M. Xue, and C. Liu, 2018: Development of a Hybrid En3DVar Data Assimilation System and Comparisons with 3DVar and EnKF for Radar Data Assimilation with Observing System Simulation Experiments. *Mon. Wea. Rev.*, **146**, 175-198.

Kong, R., M. Xue, A. O. Fierro, Y. S. Jung, C. S. Liu, E. R. Mansell, and D. R. MacGorman, 2020: Assimilation of GOES-R Geostationary Lightning Mapper Flash Extent Density Data in GSI EnKF for the Analysis and Short-Term Forecast of a Mesoscale Convective System. *Mon. Wea. Rev.*, **148**, 2111-2133.

Kotsuki, S., Y. Ota, and T. Miyoshi, 2017: Adaptive covariance relaxation methods for ensemble data assimilation: experiments in the real atmosphere. *Quart. J. Roy. Met. Soc.*, **143**, 2001-2015.

Liu, C., and M. Xue, 2016: Relationships among Four-Dimensional Hybrid Ensemble-Variational Data Assimilation Algorithms with Full and Approximate Ensemble Covariance Localization. *Mon. Wea. Rev.*, **144**, 591-606.

682 Liu, C., M. Xue, and R. Kong, 2019: Direct Assimilation of Radar Reflectivity Data Using
683   3DVAR: Treatment of Hydrometeor Background Errors and OSSE Tests. *Mon.*
684   *Wea. Rev.*, **147**, 17-29.
685 Lojou, J.-Y., and K. Cummins, 2005: On the representation of two- and three-dimensional
686   total lightning information. *85th AMS Annual Meeting*.
687 Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP - a comparison
688   with 4D-Var. *Quart. J. Roy. Met. Soc.*, **129**, 3183-3203.
689 Lu, X., X. G. Wang, Y. Z. Li, M. J. Tong, and X. L. Ma, 2017: GSI-based ensemble-
690   variational hybrid data assimilation for HWRF for hurricane initialization and
691   prediction: impact of various error covariances for airborne radar observation
692   assimilation. *Quart. J. Roy. Met. Soc.*, **143**, 223-239.
693 MacGorman, D. R., and Coauthors, 2008: TELEX The Thunderstorm Electrification and
694   Lightning Experiment. *Bull. Amer. Meteor. Soc.*, **89**, 997-1013.
695 Maldonado, P., J. Ruiz, and C. Saulo, 2020: Parameter Sensitivity of the WRF-LETKF
696   System for Assimilation of Radar Observations: Imperfect-Model Observing
697   System Simulation Experiments. *Wea. Forecasting*, **35**, 1345-1362.
698 Mansell, E. R., 2014: Storm-Scale Ensemble Kalman Filter Assimilation of Total
699   Lightning Flash-Extent Data. *Mon. Wea. Rev.*, **142**, 3683-3695.
700 Mansell, E. R., D. R. MacGorman, C. L. Ziegler, and J. M. Straka, 2002: Simulated three-
701   dimensional branched lightning in a numerical thunderstorm model. *J. Geophys.*
702   *Res. Atmos.*, **107**.
703 Mansell, E. R., D. R. MacGorman, C. L. Ziegler, and J. M. Straka, 2005: Charge structure
704   and lightning sensitivity in a simulated multicell thunderstorm. *J. Geophys. Res.*
705   *Atmos.*, **110**.
706 Marchand, M. R., and H. E. Fuelberg, 2014: Assimilation of Lightning Data Using a
707   Nudging Method Involving Low-Level Warming. *Mon. Wea. Rev.*, **142**, 4850-
708   4871.
709 Mason, I. B., 2003: Binary events. *Forecast Verification-A Practitioner's Guide in*
710   *Atmospheric Science*, 37-76.
711 Pan, Y., M. Xue, K. Zhu, and M. Wang, 2018: A prototype regional GSI-based EnKF-
712   variational hybrid data assimilation system for the Rapid Refresh forecasting
713   system: Dual-resolution implementation and testing results. . *Adv. Atmos. Sci.*, **35**,
714   518-530.
715 Pan, Y., and Coauthors, 2014: A GSI-Based Coupled EnSRF-En3DVar Hybrid Data
716   Assimilation System for the Operational Rapid Refresh Model: Tests at a Reduced
717   Resolution. *Mon. Wea. Rev.*, **142**, 3756-3780.
718 Pu, Z. X., S. X. Zhang, M. J. Tong, and V. Tallapragada, 2016: Influence of the Self-
719   Consistent Regional Ensemble Background Error Covariance on Hurricane Inner-
720   Core Data Assimilation with the GSI-Based Hybrid System for HWRF. *J. Atmos.*
721   *Sci.*, **73**, 4911-4925.
722 Qie, X. S., R. P. Zhu, T. Yuan, X. K. Wu, W. L. Li, and D. X. Liu, 2014: Application of
723   total-lightning data assimilation in a mesoscale convective system based on the
724   WRF model. *Atmos. Res.*, **145**, 255-266.
725 Rison, W., R. J. Thomas, P. R. Krehbiel, T. Hamlin, and J. Harlin, 1999: A GPS-based
726   three-dimensional lightning mapping system: Initial observations in central New
727   Mexico. *Geophys. Res. Lett.*, **26**, 3573-3576.

Skamarock, W. C., and Coauthors, 2008: A DESCRIPTION OF THE ADVANCED RESEARCH WRF VERSION 3. *NCAR Technical Note NCAR/TN-475+STR, doi:10.5065/D68S4MVH.*

Snook, N., M. Xue, and Y. Jung, 2012: Ensemble Probabilistic Forecasts of a Tornadic Mesoscale Convective System from Ensemble Kalman Filter Analyses Using WSR-88D and CASA Radar Data. *Mon. Wea. Rev.*, **140**, 2126-2146.

Sobash, R. A., G. S. Romine, C. S. Schwartz, D. J. Gagne, and M. L. Weisman, 2016: Explicit Forecasts of Low-Level Rotation from Convection-Allowing Models for Next-Day Tornado Prediction. *Wea. Forecasting*, **31**, 1591-1614.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Mon. Wea. Rev.*, **136**, 5095-5115.

Tong, M., and M. Xue, 2008: Simultaneous estimation of microphysical parameters and atmospheric state with simulated radar data and ensemble square root Kalman filter. Part I: Sensitivity analysis and parameter identifiability. *Mon. Wea. Rev.*, **136**, 1630-1648.

Wang, Y., J. Gao, P. S. Skinner, K. Knopfmeier, T. Jones, P. L. Heiselman, and L. J. Wicker, 2018: Test of a weather-adaptive dual-resolution Hybrid 3DEnVAR and WRF-DART analysis and forecast system for severe weather events. *22nd Conf. on Integrated Observing and Assimilation Systems for the Atmosphere, Ocean, and Land Surface (IOAS-AOLS)*, Austin, TX, Amer. Meteor. Soc., 168, https://ams.confex.com/ams/198Annual/webprogram/Paper330920.html.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913-1924.

Whitaker, J. S., and T. M. Hamill, 2012: Evaluating Methods to Account for System Errors in Ensemble Data Assimilation. *Mon. Wea. Rev.*, **140**, 3078-3089.

Wu, W., D. F. Parrish, E. Rogers, and Y. Lin, 2017: Regional Ensemble-Variational Data Assimilation Using Global Ensemble Forecasts. *Wea. Forecasting*, **32**, 83-96.

Xiao, X., J. Sun, X. Qie, Z. Ying, L. Ji, M. Chen, and L. Zhang, 2021: Lightning Data Assimilation Scheme in a 4DVAR System and Its Impact on Very Short-Term Convective Forecasting. *Mon. Wea. Rev.*, **149**, 353–373.