EGU22-4254
https://doi.org/10.5194/egusphere-egu22-4254
EGU General Assembly 2022

An observation operator for geostationary lightning imager data assimilation in storm-scale numerical weather prediction systems

Pauline Combarnous, Felix Erdmann, Olivier Caumont, Eric Defer, and Maud Martet
Centre National de Recherches Météorologiques, France (pauline.combarnous@mailbox.org)

In spite of the continuous improvement of numerical weather prediction (NWP) systems, thunderstorms remain hard to predict with accuracy. This difficulty partly results from a lack of observations to describe the initial state of the atmosphere. Total lightning is a good indicator of convective activity and lightning data assimilation could improve the prediction of thunderstorms, especially in regions where storm-related observations are scarce.

The Lightning Imager (LI) onboard the Meteosat Third Generation (MTG) satellites will provide total lightning observations continuously over Europe with a spatial resolution of a few kilometers. This makes it a rich potential data source for convection-permitting NWP.

To prepare the assimilation of the flash extent accumulation (FEA) measured by LI in the French storm-scale regional AROME NWP system, a lightning observation operator is required to convert the model variables into a product comparable to the observations. Since LI FEA observations are not available yet (launch planned for the end of 2022), pseudo-LI FEA observations are generated from the records of the Météorage VLF ground-based lightning detection system (Erdmann et al., 2021).

Since neither flashes nor the electric field are predicted by the AROME model, the observation operator relies on proxy variables to link the flash observations to the prognostic variables of the model. This study focuses on the evaluation of different FEA observation operators from various proxies encountered in the literature and calculated from the outputs of 1 h AROME-France forecasts for 47 electrically active days in 2018.

Different regression techniques, linear regression as well as machine learning models, are used to relate the synthetic FEA and the modeled proxies. The data are processed as distributions over the whole domain (i.e. France) and time period since a pixel-to-pixel comparison exhibits a rather poor correlation. The training of the observation operator is performed on 44 days of the dataset and 3 days are used for the validation. The performances of each observation operator are evaluated by computing Fraction Skill Scores between synthetic FEA and proxy-based FEA. Two different proxy types emerged from the literature review: microphysical and dynamical proxies. The present study suggests that microphysical proxies seem to be more suited than the dynamical ones to model satellite lightning observations.

The performances of multivariate regression models are also evaluated by combining several proxies after a feature selection based on a principal component analysis and a proxy correlation study.