Soil dust aerosols created by wind erosion of arid and semi-arid surfaces are amongst the most significant contributors to the global aerosol mass load and dominate climate effects over large areas of the Earth. They are mixtures of different minerals, whose relative abundances, particle size distribution (PSD), shape, surface topography, and mixing state influence their effect upon climate. However, Earth System and Chemical Transport Models typically assume that dust aerosols have a globally uniform composition, neglecting the known regional variations in the mineralogy of the sources. The omission of mineralogy impedes further understanding of the dust role in the Earth system. An on-going ERC Consolidator Grant called FRAGMENT (FRontiers in dust minerAloGical coMposition and its Effects upoN climaTe) was designed to fill this gap. FRAGMENT combines field campaigns, new theory, remote spectroscopy, and modeling to understand and constrain the global mineralogical composition of dust along with its effects upon climate. In this context, the eFRAGMENT1 project, awarded through the 17th PRACE Call for Project Access, is designed to tackle the modeling and HPC related activities of FRAGMENT during the 1st year of the project. One of the paths followed was to minimize the errors in the spatial and temporal variability of the simulated dust emission fluxes through the assimilation of satellite-derived dust optical depth.

We have focused on the design of appropriate ensemble members to represent the model uncertainty. As a first step, we have run and evaluated high-resolution simulations at regional scale using the Multiscale Online Non-hydrostatic AtmospheRe CHemistry model (MONARCH) coupled to a Local Ensemble Transform Kalman Filter (LETKF) data assimilation system, both developed at the Earth Sciences Department of the BSC, with dust optical depth observations from MODIS Deep Blue at 10 km resolution. To perform the simulations, we developed an automated tool to manage and monitor runs in the MareNostrum IV supercomputer of the Barcelona Supercomputing Center based on the Autosubmit workflow manager. The workflow was optimized using wrappers feature, to pack model ensemble runs and data assimilation system in a unique job of more than 19000 cores. Using this system, we have run 12 ensemble members for a domain covering Northern Africa, Middle East and Europe at 10 km resolution. We have run a total of 16 years, where 10 of them have been used for testing different approaches and for calibrating the ensemble members. With the selected configuration, we have produced six years (2011-2016) of 3 hourly output, which means more than 200T of data that is currently being evaluated with a variety of observations from ground-based and satellite observations. Overall the statistical scores of the analysis are satisfactory and superior to the control and first guess simulations.