The article "Deciphering Organization of GOES–16 Green Cumulus (Cu), through the EOF Lens" by Tom Dror et al. applies an Empirical Orthogonal Function (EOF) analysis to so-called green cumulus clouds, i.e., cumulus clouds observed over continental forested or vegetated areas, observed through the GOES-16 satellite dataset. The article is basically a proof of concept, that a carefully-done EOF analysis can reveal a number of well-documented patterns in such 2+1 dimensional data. EOF analysis is a widely-used tool in climate and atmospheric science, but, as the authors point out, is typically applied to much larger scales and usually much more poorly-resolved (both in space and time) data. The data at hand, the GOES-16 datasets, is special in that its spatial (.5 km) and temporal (5 minutes) resolution are comparably high.

The EOF analysis applied is fairly standard, but carefully carried out: the data (channel 2, 0.64 micrometers, which reflects visible solar radiation) is used for daytime detection of Cu. The data are first converted to reflectance using a referenced technique and are processed by a gamma correction (I suppose, to be able to better view the data with the bare eye?), and re-projected onto a lat-lon grid. They then focus on several subregions of the data within Eastern North America (including a section of the Appalachian Mountains) during a 2018 summer period.

The EOF analysis is standard, in that the temporal fluctuations are first computed and then the first 20 leading eigenvalues and corresponding eigenvectors of the anomaly matrix are computed. The authors perform standard statistical tests to assess the sampling error. The first 20 EOFs are found to account for roughly 2/3 of the variance.

Already without discussing the EOFs, preliminary analysis reveals a clear diurnal cycle, which is not surprising given the continental surface. The temporal mean of the reflectance field already shows cloud streets, which are fairly stationary in time. Focusing on two sub-regions, the authors further discuss the presence of a gravity wave, which appears to modify the local cloud size and structure. To assess the organization of the cloud field they compute the I$_{org}$ index, an index that essentially assesses the departure of the centroid-centroid distance distribution from random (Poisson distribution). The authors, perhaps as expected, find more regular (lattice-like) organization than random for both sub-regions.
They then turn to the EOF decomposition and use the EOFs, to further explain the pattern already discussed qualitatively. EOF1 is associated with cloud streets and the corresponding eigenvalue has a time dependence, where a peak is found around noon, hence an indication that cloud streets build up during the day under diurnal heating, but dissipate at night. The second principal component is mentioned to also show a diurnal cycle, but is later discussed as stemming from orographic clouds. The third principal component is considered to be associated with cirrus clouds. They then proceed to discuss the time dependence of all principal components analyzed (1-20) and find that, given the error bars, several of the EOFs should be seen as degenerate. They point out that the higher-frequency modes (EOF10-EOF20) are able to capture the period of individual clouds. They finally show that the actual field can be well reconstructed from the first 20 modes they analyzed.

The authors point out that the study is essentially a proof of concept and that it allows decomposition of the cloud field into modes of different frequencies and explicit detection of cloud streets, orographic clouds and gravity wave effects.