I welcome this paper since it bravely deviates from the usual and common implementation of Optimal Estimation (OE). In doing so, the paper largely agrees with retrieval strategies so long pursued within IASI activities by my group and colleagues of mine.

Having said that, I have to stress that the fact that OCS can be retrieved with a poor a prior constraint and based only on IASI data has been recently put forward by Liuzzi et al 2016 (doi:10.1016/j.jqsrt.2016.05.022), who demonstrates and exemplifies how OCS column amount can be retrieved with a degree of freedom of 1 and a precision for the single IASI field of view of 6-7%. In this respect, the paper by Vincent and Dudhia is lacking important references and should be updated.

The "unprecedented features" (pag, 2 lines 7-17) claimed in the paper are common practice within the IASI community.

The use of *ensemble of states* to initialize the non-linear retrieval scheme dates back to Chedin et al, J. *Climate Appl. Meteor.,24, 128-143, 1985* (TIGR, Thermodynamic Initial Guess Retrieval, http://ara.abct.lmd.polytechnique.fr/index.php?page=tigr). Furthermore, the approach has been used in the context of OE to linearize the radiative transfer equation by (to name a few) Grieco et al 2007 (doi:10.1002/qj.162), Masiello et al 2009 (doi:10.5194/acp-9-8771-2009) who used an ensemble of states to train an EOF regression for the purpose of computing a suitable background around which to linearize the inverse problem, instead of simply using the *ensemble-mean atmosphere*. This approach, i.e. EOF initialization, would greatly improve the algorithm shown in Vincent and Dudhia paper, e.g., to improve the linearity in H_2O channels. Presently, Liuzzi et al 2016 linearizes the OE retrieval with time-space collocated ECMWF analysis.

The use of a *data driven retrieval approach* which simultaneous (*jointly*) retrieves temperature and gases dates back to Smith et al, 1991 (http://dx.doi.org/10.1364/AO.30.001117) and has been extended by Liuzzi et al 2016 to simultaneously (jointly) retrieve emissivity, surface temperature, H_2O, HDO, O_3, CO_2, N_2O, CO, CH_4, CF_4, SO_2, NH_3, HNO_3, and finally OCS. Previous specific application to gas retrieval alone can be found in Grieco et al 2013 (doi:10.1364/OE.21.024753).

As I said, this paper provides an interesting approach to the design and implementation of a non-conventional OE data driven retrieval strategy, which is specifically designed with a loose a-prior constraint. In this respect, authors should be encouraged. However, they should correctly acknowledge work already done in the same area. In addition, there are many weak points that should be properly addressed.

The authors fails to show how the retrieval products are correlated, one would rather expect a strong correlation between OCS and T_s since the authors seem to privilege atmospheric window channels. Beyond 2000 cm^{-1}; radiances (especially window channels) are contaminated by solar radiation. Authors avoid sunglint, however solar contamination is in any case to be taken into account in daytime. Over land, especially for desert, arid and semi-arid regions, the emissivity is strongly dependent on the wavenumber, so that the use of an effective εT_s is not adequate. The CO_2 spike at 2077 cm^{-1} (Fig. 5 in the paper) is not seen in state-of-art forward modelling calculations (e.g., once again see Liuzzi et al 2016). Therefore, how accurate is authors' RFM? It is not clear how the radiative transfer is manipulated to arrive at the state vector defined in Eq. (7). The state parameters seem to have been included rather ad hoc, whereas the state vector should be derived by a coherent linearization of the radiative transfer equation, accounting for the physical parameters that contribute to the radiance signal and showing how these can be transformed to the state vector of Eq. (7). Finally, what about CO and O_3 retrievals? Why they are not shown.