IRS2012-41-2
International Radiation Symposium 2012
Dahlem Cube, Berlin, Germany, 06 – 10 August 2012

An extended data validation test for parametrized retrieval algorithms

M. Schaale (1) and Th. Schroeder (2)

(1) Freie Universitaet Berlin, Institute for Space Sciences, Department of Earth Sciences, Berlin, Germany (michael.schaale@fu-berlin.de), (2) CSIRO Land and Water, Environmental Earth Observation Group, Brisbane, Australia (thomas.schroeder@csiro.au)

The retrieval of environmental data from multi-spectral remotely sensed data is very often based on the (partial) inversion of extensive radiative transfer simulations (RTS). The inversion can be utilized in different ways, e.g. through the usage of polynoms or artificial neural networks. The inversion algorithms (IA) usually contain numerous parameters which have to be adapted by regression schemes in a training phase with the help of the RTS data. The subsequent processing of real remotely sensed data by an adapted IA requires a validity test (VT) of the input data (usually a vector consisting of TOA radiances, environmental and geometric data) before feeding them into the IA. This ensures that these or similar data were included in the training phase of the IA and helps to avoid unpredictable extrapolation effects. In standard procedures these 'out-of-scope' data are identified by a simple convexity test (CT). CT means that each dimension of the input vector is tested to lie between the minimum and maximum of the corresponding dimension of the used training data set. This assumption is rather crude as it assumes a homogeneously filled data space. But in general the data are not distributed homogeneously and thus a CT is an incomplete and unsatisfactory check.
This paper proposes a solution to the problem sketched above by the development and implementation of an enhanced VT (eVT) which is based on a density map of the data space. The density map itself is approximated by an extended neuronal vector quantization method. The newly developed eVT algorithm is tested with known distributions of artificial data.
Although the eVT is not limited to a specific retrieval/inversion scheme it is finally applied to an existing retrieval scheme for coastal water constituents from satellite data (MERIS) acquired over coastal regions in Europe (here:FUB/WeW water processor for VISAT-BEAM). A comparison against the data filtered by a simple CT further stresses the importance of a thorough data test.