Introduction to the Validation Module Design for CMDPS Baseline Products

Shin-Young Kim, Chu-Yong Chung, and Mi-Lim Ou

Global Environmental System Research Laboratory, National Institute of Meteorological Research, KMA
Seoul 156-720, solarh@metri.re.kr

ABSTRACT

CMDPS (COMS Meteorological Data Processing System) is the operational meteorological products extraction system for data observed from COMS (Communication, Ocean and Meteorological Satellite) meteorological imager. CMDPS baseline products consist of 16 parameters including cloud information, water vapor products, surface information, environmental products and atmospheric motion vector. Additionally, CMDPS includes the function of calibration monitoring, and validation mechanism of the baseline products. The main objective of CMDPS validation module development is near-real time monitoring for the accuracy and reliability of the whole CMDPS products. Also, its long time validation statistics are used for upgrade of CMDPS such as algorithm parameter tuning and retrieval algorithm modification. This paper introduces the preliminary design on CMDPS validation module.

Key words: COMS, validation, collocation, statistics

1. INTRODUCTION

Communication, Ocean and Meteorological Satellite (COMS), planned for launch in 2009, is the first Korean multi-purpose geostationary satellite with the three main missions (communication, ocean, and meteorological applications). For the meteorological purpose, COMS will be mainly used for real-time monitoring weather phenomena, providing improved input data for numerical weather prediction models, and monitoring climate change (METRI, 2004). With the development of the COMS system, CMDPS (COMS Meteorological Data Processing System) is focused to utilize the COMS observed data. The primary function of CMDPS is to derive 16 level 2 environmental products using the geo-located and calibrated level 1B data. Algorithms and integrated processing system for 16 baseline products have been developed. The baseline products include cloud information, cloud type, rain rate, sea/land surface temperature and emissivity, water vapor information, insolation, and so on. Figure. 1 is Hierachical structure of CMDPS system components from the system level to the program level. CMDPS includes pre-processing and post-processing modules, data processing modules, and off line modules such as the calibration and validation (METRI, 2005). To operate the Validation module (VAM), we need to accumulate the collocation data set of the reference data and CMDPS's 16 baseline parameters, then get statistics of the accumulated data set. VAM is the module which performs verification according to validation criteria and provides a result, using validation method, collocation criteria and the reference data (METRI, 2007).

In order to validate COMS data products, it is necessary to validate specific atmosphere parameters under a wide variety of atmospheric conditions, solar illumination angles, and ecosystems worldwide.

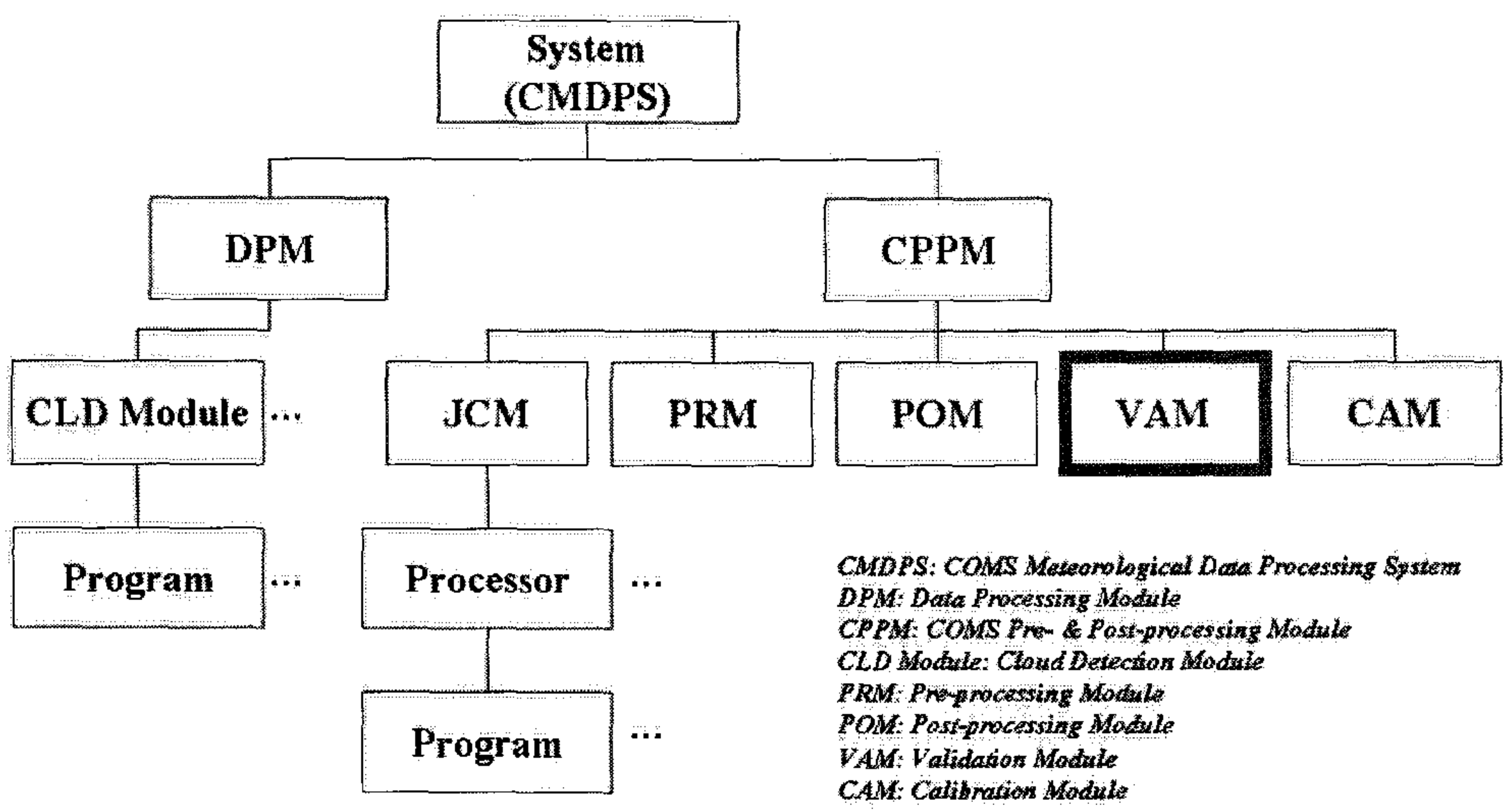

Figure. 1. Hierachical structure of CMDPS system component, from the system level to the program level.

2. VALIDATION CRITERIA

CMDPS will use various validation methods and reference data to assess the accuracy of its 16 baseline products (Table1). The reference data include (i) in situ data collected from over a distributed network of ground observation sites (e.g. radiosonde, AWS, buoy, etc.), (ii) data and products from other satellite sensors (e.g. MODIS, CERES, SSM/I and TRMM) and (iii) data from radiative transfer simulation.

Table 1. The reference data to assess the accuracy of 16 - baseline products.

Standard data		CMDPS products
Ground_ Observation	Ground	INS, FOG, CA
	AWS	LST, RI
	BUOY	SST
Upper level_ Observation	radiosonde	CSR, TPW, UTH, AMV
	wind profiler	AMV
Satellite_ Products	MODIS	CLD, SSI, LST, CT, CTTH, COD, CP, AI, AOD
	SSMI	SSI, RI
	CERES	OLR
	OMI	AI, AOD
	SATOB	AMV
NWP		CSR, AMV

Ground based observation data which are measured every hour over global area are used to validate the parameter such as Insolation (INS), Fog detection (FOG), Cloud Amount (CA). And AWS data are used to validate Land Surface Temperature (LST) and Rain Intensity (RI). Sea Surface Temperature (SST) is used for comparison with BUOY drafter data which is produced for the 6–hour interval. CMDPS also plan to utilize ground-based radiometer observations to derive Total Precipitable Water (TPW), Upper Tropospheric Humidity (UTH) and Atmospheric Motion Vector (AMV), especially over the global site. Sonde sounding data, on the utilization of four times a day (00, 06, 12 and 18UTC) are collected from Global Telecommunication System (GTS). The comparison between COMS data and sonde measurements was performed for coincident profiles. CMDPS cloud properties will be intercompared with those derived from MODIS level 2 products (Cloud Detection (CLD), Cloud Phase (CP), Cloud Optical Thickness (COT), Cloud Fraction (CF), Cloud Top Temperature and Cloud Top Height (CTTH)) and ground observation data. Also, Aerosol optical depth (AOD) and Aerosol Index (AI) from CMDPS will be compared with MODIS aerosol retrieval data as well as to TOMS/OMI (Total Ozone Mapping Spectrometer/the Ozone Monitoring Instrument) aerosol index. Sea Ice/Snow cover detection (SSI) will be compared with (i) MODIS level 3 SSI data, (ii) IMS measurements over the north hemisphere, and (iii) SSM/I data. Atmospheric motion vector will be compared with (i) wind profiler, (ii) radiosonde, (iii) NWP data, and (iv) SATOB measurements.

CMDPS validation work will involve collocation in space and time with the reference data. CMDPS data and the reference satellite data (MODIS, CERES, etc) differ in spatial and temporal resolution. Pixel-by-pixel comparison with collocated satellite observation attains quantitative validation results. In the case of ground station data (ground, AWS, and buoy) and upper level observation (radiosonde, wind profiler), area average of CMDPS products around the observation station is used in calculation. Also a key component of the validation exercise, in addition to the collection of data and comparison with CMDPS retrievals, is the concomitant increase in understanding the operation of the CMDPS system (sensor plus algorithms) required to fine-tune the algorithms in a rational manner. This fine tuning process is particularly important that the algorithms for 16 baseline parameters provide consistent and accurate results.

3. STATISTICAL METHODS FOR VALIDATION

Validation methods are determined in consideration of the target variable types (continuous or categorical). For validation of continuous type variables (LST, COT, INS, etc), simple mean, bias and the correlation coefficient are used for quantifying the prediction ability of corresponding variables. On the other hands, categorical statistics quantify skill in the prediction of the occurrence of an event. Categorical statistics are needed to evaluated binary, yes/no, CMDPS data of the type of statements that an event will or will not happen. The first step to verify binary data is to compile a 2x2 contingency tables showing the frequency of yes-no-type CMDPS data and corresponding reference data. These measures include Percent Correct (PC), Probability of Detection (POD), False Alarm Ratio (FAR), Threat Score (TS), etc.

In case of binary variables, the representative parameters depend on the occurrence of the event. For example, if fog occurs frequently in the validation domain, one can use PC, PSS, and HSS as the statistical parameters, in other case, one can use CSI and GSS.

Also, categorical events are naturally not limited to binary data of two categories and the associated 2x2 contingency tables. The distributions approach to forecast verification examines the relationship among the elements in multi-category contingency tables. One can consider local weather variables in several mutually exhaustive categories; for example cloud type is classified into Cirrus/Cirrostratus/Deep-convection/Alto-cumulus/Altostratus/Nibostratus/Cumulus/Stratocumulus /Stratus (k=9) and likewise for cloud phase categorized into ice/water/mixed/uncertained (k=4). The PC, KSS and HSS skill scores can be generalized to multi-category cases (Nurmi, 2003). Both KSS and HSS are measures of potential improvement in the number of correct forecasts over random forecasts. These values show the accuracy of multi-categorical variables on the whole.

Table 2. Statistical parameters used for estimation of the CMDPS product accuracy.

Product	Variable's type	Statistical parameter
CLD	binary variable	PC, PSS, HSS, POD, FAR
SSI	binary variable	POD, FAR
SST	continuous variable	bias, rmse, r
LST	continuous variable	bias, rmse, r
INS	continuous variable	bias, rmse, r
TPW	continuous variable	bias, rmse, r
UTH	continuous variable	bias, rmse, r
CT	multi-variable case	PC, PSS, HSS
CA	continuous variable	bias, rmse, r
CTHH	continuous variable	bias, rmse, r
COT	continuous variable	bias, rmse, r
CP	multi-variable case	PC, PSS, HSS
Fog	binary variable	PC, PSS, HSS, CSI, GSS
RI	continuous variable	bias, rmse, r
AI	binary variable	PC, PSS, HSS, POD, FAR
AOD	continuous variable	bias, rmse,r
OLR	continuous variable	bias, rmse, r
AMV	continuous variable	bias, rmse, r

4. SUMMARY

The primary validation procedure includes reference data collection, comparison and collocation according to time and space. CMDPS plans to gather all relevant data from various validation exercises and make these data available in timely manner. At the results, CMDPS which produces 16 baseline products will run the validation module regularly to produce validation results on operational basis. Also, its long time validation statistics will be feedback to update the CMDPS algorithm.

REFERENCES

Nurmi, P., 2003. Recommendations on the verification of local weather forecasts. ECMWF Technical Report No 430

METRI/KMA, 2004. Development of Meteorological Data Processing System of Communication, Ocean, and Meteorological Satellite(I), pp. 421.

METRI/KMA, 2005. Development of Meteorological Data Processing System of Communication, Ocean, and Meteorological Satellite(II), pp. 846.

METRI/KMA, 2007. Development of Meteorological Data Processing System of Communication, Ocean, and Meteorological Satellite(IV), pp. 618.

Acknowledgements

This study has been performed as a part of "Development of Meteorological Data Processing System for COMS" project at the Korea Meteorological Administration.