CGMS-51-KMA-WP-03
28 June 2023

Prepared by: KMA
Agenda Item XX
Discussed at Plenary

Subject	**KMA PLANS FOR AI/ML EXPLOITATION**
In response to CGMS action/recommendation	
HLPP reference	
Executive Summary	Currently, NMSC/KMA intends to develop a strategic action plan, building on existing experience and facilitating the exploitation of satellite data including GK2A etc. through the AI/ML potential, understanding its limitation and strengths. NMSC/KMA is considering 4 development perspectives for AL/ML. And we made a roadmap from 2022 to 2024 which contains experimental and feasibility research projects. This may be pre-phase of GK5(follow-on of GK2A) products development. In this paper, our challenge AI products such as proxy data etc. are introduced. Also, this paper introduces KMA Progress on “cloud-based data hub” to enhance the accessibility and usage of meteorological and climate data including for external users. KMA are also making “Cloud based data hub system” to enhance accessibility and usage of various meteorological and climate data including satellite data for external users. It is a consumer-customized APIs, established as an open platform that can create added value by supporting the decision-making of the country and society by integrating and providing links based on Climatology and it is also data analysis platform for external users applies cloud technology to provide users with computing resources, meteorological and climate data, software (R, Python, etc.) required for data analysis, and visualization analysis functions.
Action/Recommendation proposed	

1 KMA PLANS FOR AI/ML EXPLOITATION

1.1 AI/ML APPLICATION FOR SUPPORTING WEATHER AND CLIMATE AT KMA

Currently, NMSC/KMA intends to develop a strategic action plan, building on existing experience and facilitating the exploitation of satellite data including GK2A etc. through the AI/ML potential, understanding its limitation and strengths. AI/ML has proven its potential for satellite in key action that could be good starting point to identify priorities for NMSC/KMA roadmap. NMSC/KMA has 4 development perspectives for AL/ML. First one is to extract the blended products considering meteorological radar, satellite data and in-situ measurement for feature selection and extraction. Secondary one is to automatically extract extreme pre-condition or forecast mesoscale system life cycle with hazardous weathers. Third one is to extract higher quality grid products for in-situ or ground radar absent area by using satellite data. Forth one is to improve decision making processes by making super-resolution information data through AI technique. These perspectives can be modified in future.

We made a roadmap from 2022 to 2024 which contains experimental and feasibility research projects. This may be pre-phase of GK5(follow-on of GK2A) products development.

Especially, we found AI products has better accuracies than those of conventional algorithm. High quality AI products make important role in evaluating qualities of in-situ measurements and reducing budget for their maintenance. We will expand AI technique to more in-situ measurements including GHG (CO2) and estimate long-term AI based products (EVT, SM, LST, SST etc.) for supporting climate mission such as drought or heatwave over Korean region.

Meanwhile, KMA has provided true colour RGB images and day-night RGB images at night-time as well as daytime in public since 24th March, 2021. The GK2A proxy visible data are generated based on one of AI-based technique, Conditional Generative Adversarial Networks (CGAN). The main characteristics of CGAN is to make model learning by competition of two kinds of objects which can generate accurate results. The GK2A AI-based true colour RGB image can identify cloud and surface type in night-time, which was impossible with original true colour RGB image. Also, AI-based day-night RGB images which are basically used for weather status analysis by forecasters can identify low-level cloud with distinct red colour, while it is difficult with original day-night RGB image which has no night-time visible data.

This is other proxy data for monitoring heavy rain through AI technique, we have no radar observation in western and southern part of Korea which is very significant area for monitoring severe weather such as typhoon or developed convective clouds crossing this area. Proxy radar is estimated from GK2A data by using AI technique (Pix2Pix (CGAN)) for no ground radar observation and provided to forecaster every

10 minutes. Validation Results with ground radar reflectivity (June, 2022). R: 0.6/RMSE:6.0 dBz/Bias:0.3Dbz. Proxy radar tends to overestimate in terms of presence or absence of precipitation, therefore false alarm ratio is large and underestimate in terms of intensity of reflectivity, especially for stratus cloud with heavy rain which is similar result to IR based precipitation. It still needs optimization of AI model with the more tailored dataset.

1.2 ENHANCEMENT OF USAGE OF SATELLITE DATA BY "CLOUD BASED DATA HUB SYSTEM" AT KMA

KMA has created a "weather Climate Data Hub system" (hereinafter referred to as "data hub") that can conveniently and freely analyse and utilize large-scale meteorological and climate data in the field of disaster prevention and the 4th industry at a time when climate crises such as torrential rain and typhoon."

Data Hub uses the vast amount (hundreds of petabytes) of global meteorological and climate data accumulated by KMA and related organizations Disaster sites and the 4th industry convergence (artificial intelligence, big data).

Consumer-customized APIs is established as an open platform that can create added value by supporting the decision-making of the country and society by integrating and providing links based on Climatology.

KMA plans to divide the data hub construction into four stages and complete the establishment of a backup system in preparation for system disasters such as fires by spending a total project cost of about KRW 30 billion by 2025. It is said to be in the second half of 2023.

From 2023, citizens can select meteorological and climate data through a customized API* method and receive real-time service, and government agencies, public institutions, industries, and academia in each sector of society can use data processing without building large-scale systems required for data utilization. It is expected that business processing will be possible without the necessary expertise. In response to the high demand for the use of meteorological and climate data, KMA is sharing and utilizing data to help users more conveniently collect and process the data they need. This system will be converted into a custom API method.

The number of downloads of large-capacity grid data such as numerical models, satellites and radars a significant jump from 2.5 million in 2020 to 11 million at the end of September 2022.

Development of an API applied to the data format conversion function that was provided only in the API method for observations, forecasts (neighbour-hood

forecasts, special reports, etc.) and daily weather index, gradually expanding to large-capacity grid data* such as numerical models, satellites, and radars. policy.

In addition, the API method will be configured in a customized format that allows consumers to select and receive only the data on the desired variable or desired region among the data and receive it in real time.

KMA expects that this will enable consumers to conveniently develop web and apps without having to have a separate data storage and management system, and significantly reduce the time required for data processing, resulting in a great contribution to vitalizing data utilization.

Along with this, KMA plans to build and test-operate a cloud-based data analysis and utilization system to induce the analysis and utilization of meteorological and climate data in the private sector. The data analysis/utilization system is a data analysis platform for external users of KMA. It applies cloud technology to provide users with computing resources, meteorological and climate data, software (R, Python, etc.) required for data analysis, and visualization analysis functions.

As a result, it is expected that external users in the private sector will be able to more easily search and store various meteorological and climate data in the KMA's internal system, and perform and visualize more diverse and complete data analysis.