

클라우드 컴퓨팅 기반 사용자 맞춤형 계절 예측자료 추출 및 가공서비스 개발

한 정 민^{1*} · 임 창 목²

¹*APEC기후센터 기후분석과 선임연구원

²APEC기후센터 예측운영과 연구원

Development of user-customized service for seasonal forecast data processing and extracting based on Cloud Computing

Jeong-Min Han¹ · Chang-Mook Lim^{2*}

¹*Research Fellow, Climate Analytics Department, APEC Climate Center, Busan 48058, Korea

²Researcher, Climate Prediction Department, APEC Climate Center, Busan 48058, Korea

[요 약]

기후변화로 인해 지구 전역에서 이상기후 현상이 발생하고 있어 기후자료를 활용한 예측정보는 재난·재해를 방지하기 위해 중요한 정보로 활용되고 있지만, 작은 이상기후 현상에도 큰 피해를 겪는 동남아시아 등 아시아 태평양의 국가들은 예측정보 생산을 위한 자료수집 및 처리에 필요한 전산 자원의 부족으로 기후자료 활용도가 낮다.

그리하여, 주요 모델 생산기관으로부터 수집된 계절 예측자료를 활용하여 강수나 기온 등, 사용자가 원하는 변수와 지역에 대해 정보를 추출하여 사용자가 쉽게 이해할 수 있는 파일 포맷으로 가공하여 제공하는 서비스를 개발하였다. 이를 통해, 자료저장공간을 줄이고, 시각화된 이미지 정보 및 이해하기 쉬운 파일 포맷으로 제공하게 되어 별도의 자료처리 시스템 도움 없이 실시간으로 분석된 결과를 확인할 수 있었다.

[Abstract]

Due to climate change, abnormal climate phenomena are occurring all over the world, so forecast information using climate data is being used as important information to prevent disasters and disasters. countries have low uses of climate data due to the lack of computing resources necessary for data collection and processing for prediction information production.

Therefore, we developed a service that extracts information about variables and regions desired by users, such as precipitation and temperature, using seasonal forecast data collected from major model production organizations, and processes them into a file format that users can easily understand. This reduces data storage space and provides visualized image information and an easy-to-understand file format, so you can check the analyzed results in real time without the help of a separate data processing system.

색인어 : APEC기후센터, 기후자료, 계절예측자료, 다중모델 앙상블, 가공서비스

Keyword : APEC Climate Center, Climate data, Seasonal forecast data, Multi-model ensemble, Data processing

<http://dx.doi.org/10.9728/dcs.2021.22.10.1543>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 02 September 2021; **Revised** 27 September 2021

Accepted 27 September 2021

***Corresponding Author; Jeong-Min Han**

Tel: +82-51-745-3989

E-mail: goal@apcc21.org

I. 서론

기상기후 자료는 기상위성, 레이더 등의 관측 장비의 발달과 함께, 전 지구적인 온난화 및 기후변화를 이해하고 각종 기상이변을 감시하기 위한 중요한 역할을 담당하고 있다. 대기, 해양, 육지 및 생물이 상호작용하는 전 지구환경 속에서 일어나는 기후변화와 기상이변을 예측하고 현상을 이해하기 위해서는 연구자들을 위한 지속적인 자료제공으로 안정적인 기후자료서비스가 제공되어야 한다[1].

최근 들어, 기후변화로 인해 지구 전역에서 이상기후 현상이 발생하고 있다. 특히, 작은 이상기후 현상에도 동남아시아 등 아시아 태평양의 국가들은 큰 피해를 입고 있으나 예측정보 생산을 위한 전산 자원의 부족으로 재난 재해에 제대로 대비를 하지 못하고 있다.

기후 예측을 위해서는 많은 관측자료 수집 및 관리, 예측 모델 보유, 각 기관의 독자 모델 공유, 자료처리에 필요한 전산 자원 보유가 필요하다. 하지만, 동남아시아 지역이나 태평양 도서국의 경우, 기후자료를 활용한 예측정보는 재난·재해를 방지하기 위해 중요한 정보로 활용된다는 것을 알고 있지만, 예측정보 생산을 위한 자료수집 및 처리에 필요한 전산 자원의 부족으로 기후자료 활용도가 다른 국가들에 비해 낮다. 더불어, 기후자료는 일반적인 파일 포맷이 아닌, 시·공간 정보와 다양한 변수가 하나의 파일로 구성되어 있고, 대용량 정보를 압축한 바이너리 형태로 구성되어 있어 일반 연구자들이 원시 자료를 이용해 예측정보를 생산하기는 까다롭다. 또한, 대용량 파일에 전 지구 정보를 포함하고 있어 한 국가나 특정 지역만 추출하여 분석하기를 원하는 사용자에게는 자료추출 및 처리에 많은 시간이 소요된다.

이런 문제를 개선하기 위해, 전산 자원이 부족한 아시아 태평양 지역의 여러 나라뿐만 아니라 대용량 전산 자원이 구성되어 있지 않은 국내 연구자를 위해 APEC기후센터에서 운영하는 고성능 계산자원을 이용하여 자료를 추출하고 분석하여 계산된 결과(이미지, 경량화된 파일)를 인터넷을 통해 제공할 수 있는 시스템을 개발하였다.

II. 연구방법

2-1 문제 제기

미국해양대기청(NOAA)에서는 기후, 기상, 해양, 위성 등 다양한 자료를 제공하고 있으나, 사용자가 원하는 변수와 기간에 따른 자료추출 서비스를 제공하지 않고 있어 사용자는 다양한 종류의 자료에서 필요한 자료만 내려받아 개인 전산 자원을 활용하여 필요한 정보를 생산해야 한다[2]. 하지만, 대부분 자료는 미국을 중심으로 구성되어 있고 유럽의(ECMWF) 주요국 가도 상황은 미국과 같다[3].

APEC기후센터에서 제공하는 기후자료처리시스템(CLIPs)

은 자신이 원하는 자료를 추출하여 제공하는 장점이 있으나 개인 컴퓨터에 설치하여 사용하는 소프트웨어 방식으로 제공되어 있어 자료처리를 위해서는 개인 컴퓨터의 성능이 중요하다. 또한, 논문에서 제시하는 자료계산이나 시각화 기능은 제공하지 않는다[4].

일본의 자료통합분석시스템(DIAS)은 해양, 기상, 수자원 등 다양한 자료를 수집하여 제공하고 있으나, 자료검색 및 모니터링 이미지 정보를 제공하고 있어 논문에서 제기하고자 하는 자료처리, 추출, 가공 및 시각화 정보 제공과 차이를 보인다[5].

또한, 웹 인터페이스 서비스(Open Web Processing Service)는 지리정보를 이용하여 지도기반 마스킹 정보를 제공하여 공간자료를 추출하는 서비스로 해안선이 복잡한 도서 지역에 유용하지만, 해상도가 낮은 기후자료에 활용하기는 적절치 못하고 CUI (Command User Interface)로 사용을 위해서는 전문지식이 필요하다[6].

대부분 서비스는 자신의 국가를 중심으로 자료를 수집하여 제공하고 있으며 자료처리 방식 또한, 범용적이지 못하다. 아시아 태평양 도서국에서 사용하기 위해서는 최소한의 인터넷 자원과 표준화된 웹브라우저만 있으면 자료추출 가공, 결과 제공까지 활용할 수 있는 서비스가 필요하다.

2-2 제안 서비스

모바일 및 인터넷 통신기술 발달과 웹 기술의 표준화가 진행되면서 클라우드 기반의 자료처리와 웹 기반의 서비스로 분리하여 시스템 개발이 필요하다.

첫째, 자료의 측면에서는 전 세계의 대표적인 모델 생산기관의 신뢰성 있는 모델 예측자료의 수집이 필요하다. 신뢰성 있는 모델 자료를 통해 예측 결과의 신뢰성을 향상시킬 수 있기 때문이다. 이를 위해 14개의 대표적인 계절 예측 모델 생산기관으로부터 정기적으로 개별 모델자료를 수집하고 APEC 기후센터의 다중모델 앙상블 기법(MME)을 적용하여 생산된 계절 예측자료를 대상으로 서비스를 제공한다.

둘째, 사용자 서비스 측면에서는 표준화된 웹 통신기술을 활용하여 웹브라우저만 있으면 활용이 가능한 서비스 제공이 필요하다. 모바일 플랫폼부터 다양한 전산기기의 발전으로 사용자 서비스가 확장되었지만, 안정적인 서비스 제공을 위해서는 W3C에서 제공하는 표준화된 웹 표준 규약 코드를 이용하여 인터넷이 가능한 곳에서는 어떠한 단말기를 이용하든 손쉽게 사용할 수 있는 서비스를 제공하고자 한다. 더불어 대용량 원시 자료를 기반으로 가공된 결과 파일은 경량화되고 사용자 편리성을 위해 비동기식 자료전송서비스를 적용하고자 한다.

셋째, 사용자 활용 측면에서는 그래픽사용자인터페이스(GUI) 방식으로 서비스를 제공해야 한다. 주요 사용자는 컴퓨팅 자원이나 컴퓨팅 기술이 높지 않은 일반 연구자나 재난 재해를 대비하는 정책결정자를 위한 일반 사용자 수준에서 제공되어야 한다. 그리하여 서비스는 마우스만을 활용하여 자료를 처리하고 분석하여 분석된 결과를 이미지 정보나 사용

자가 쉽게 이해할 수 있는 파일 포맷으로 변환하여 제공하도록 구축하고자 한다.

넷째, 자료처리 측면에서는 모든 자료처리 및 분석은 APEC 기후센터에서 보유 중인 고성능 계산자원을 활용하여 분산병렬 처리 형태로 시스템을 구성하고자 한다. 이런 구성을 통해 빠른 계산 결과를 사용자에게 제공할 수 있고 사용자의 전산기기의 성능에 따른 결과 제공 시간이 달라지지 않기 때문이다.

Ⅲ. 대상 자료

자료는 대만, 러시아, 미국, 영국, 이탈리아, 일본, 중국, 캐나다, 한국, 호주에 있는 계절 예측을 하는 대표기관인 14개 기후 예측 현업 및 연구기관에서 제공하는 기후 예측 모델(이하 개별모델)의 계절 예측자료와 다중모델 앙상블(Multi-Model Ensemble, MME) 기법을 적용한 계절 예측 자료를 이용하였다[7].

3-1 개별 모델 예측자료

전 지구 규모의 기후 예측정보를 제공하는 세계 대표 현업 및 연구기관의 협력을 기반으로 같은 형식의 표준화된 계절 예측자료를 매일 수집하고 있다.

표 1. APEC기후센터 다중모델 앙상블에 참여하는 개별모델

Table 1. Participating individual models for APCC MME.

Institute (Country)	Model Name	Resolution	Ensemble (F/H*)	hindcast Period
APCC (Korea)	SCoPS	T159L31	10/10	1982-2013
BCC (China)	CSM1.1m	T106L26	24/24	1991-2015
BOM (Australia)	ACCESS-S	N216L85	11/11	1990-2012
CMCC (Italy)	SPS3	1x1, L46	20/20	1993-2016
CWB (Taiwan)	TCWB1Tv1.1	T119L40	30/30	1982-2019
HMC (Russia)	SL-AV	1.125x1.40625, L28	20/10	1985-2010
JMA (Japan)	JMA/MRI-CPS2	T159L60	51/10	1979-2014
KMA (Korea)	GloSea5GC2	N216L85	42/12	1991-2016
MGO (Russia)	MGOAM-2	T42L14	10/6	1979-2004
MSC (Canada)	CanSIPsv2	T63L35	20/20	1981-2010
NASA (USA)	GEOS-S2S-2.1	0.5x0.5, L72	4/4	1981-2016
NCEP (USA)	CFSv2	T126L64	20/20	1982-2010
PNU (Korea)	CGCMv2.0	T42L18	30/30	1980-현재
UKMO (UK)	GloSea5	N216L85	42/28	1993-2016

(* F: Forecast, H: Hindcast)

계절예측 자료는 크게 실시간 예측(forecast)과 과거 기후 재현(hindcast)으로 구분되며, 기관별 기후 예측모델의 특성에 따라 예측 시간(선행 시간), hindcast 기간, 해상도, 변수, 앙상블 등의 제공 정보가 상이하다. 개별 모델별 계절 예측자료는 제공 기관에 따라 원시 자료의 형식 및 구조 또한 전부 상이하지만, 해당 계절 예측자료에 대하여 일관된 형식으로 표준화된 자료를 제공하고 있으며, 결정론적 수치에 대하여 제공한다.

3-2 다중 모델 앙상블 예측자료

개별모델의 계절 예측정보를 기반으로 여러 가지 다중모델 앙상블(이하 MME) 기법을 적용하여 예측정보를 매일 생산하고 있다. MME 예측은 참여(개별)모델 간의 불확실성을 최소화하여 단일 모델이 갖는 예측성 보다 안정적이고 향상된 예측성을 나타낸다[8].

다중앙상블 기법의 다양한 기법 중에서 결정론적(deterministic) MME 예측기법과 확률론적(probability) MME 예측기법을 제공하고 있다.

결정론적 MME 예측기법으로 각 개별 모델별로 모든 앙상블 예측 구성의 평균값을 동일한 가중치로 종합하는 방법의 SCM(Simple Composite Method) 기법이 사용되었고, 1991-2010년(20년)의 기후 평년 기간에 대해 기후 평년값을 기반으로 한 편차값에 대하여 제공한다. 확률론적 MME 예측기법으로 Gaussian fitting method(GAUS)를 활용하여 각 개별 모델별 가중치를 부여한 확률분포 상에서의 3분위(평년 대비 높음/많음, 평년 비슷, 평년 대비 낮음/적음) 범주로 예측확률을 제공하며, 이 또한 SCM과 동일한 기후 평년 기간을 평년으로 정의하고 있다[9].

3-3 계절 예측자료 특성

계절예측 자료는 4개의 기후 예측모델의 예측자료와 다중모델 앙상블(MME) 자료를 활용하였다. 원시 자료는 기상/기후 분야에서 자주 사용하는 NetCDF 형태이며, 기본적으로 결정론적 예측자료는 기온(temperature at 2m; t2m), 강수(precipitation; prec), 850hPa 기온(temperature at 850hPa; t850), 해면기압(sea level pressure; slp), 해수면온도(sea surface temperature; sst), 500hPa 지위고도(geopotential height at 500hPa; z500), 850hPa 동서바람(u-vector wind at 850hPa; u850), 850hPa 남북바람(v-vector wind at 850hPa; v850), 200hPa 동서바람(u-vector wind at 200hPa; u200), 200hPa 남북바람(v-vector wind at 200hPa; v200)의 변수에 대해서 공통적으로 제공되며, 확률론적 예측자료는 기온, 강수, 850hPa 기온, 해면기압, 해수면 온도, 500hPa 지위 고도의 6종의 변수가 제공된다.

개별모델은 2020년 7월 예측분부터 해당 모델의 모든 선

행 시간에 대하여 제공되지만, MME 예측자료의 경우 개별 모델의 최대 선행 시간에 따라 3개월 예측과 6개월 예측의 두 가지 선행 시간 형태로 구성되고, 각 선행 시간별로 MME 참여모델 구성이 상이하다. 이는 일부 모델 중 6개월 미만의 예측정보를 제공하는 개별모델이 있기 때문에 3개월 예측에는 거의 모든 개별모델이 MME에 참여 가능하나, 6개월 예측에는 6개월 이상의 예측정보를 제공하는 모델들만 MME에 참여하기 때문이다.

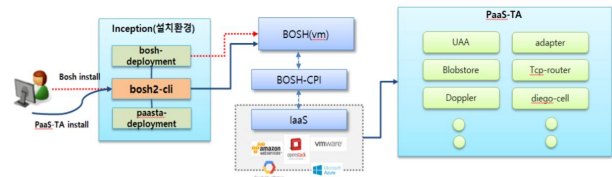


그림 1. BOSH2를 이용한 기본 시스템 기본 구성도[11]
 Fig. 1. Basic Installation Structure of PaaS-TA Using BOSH2[11]

IV. 시스템 구성

4-1 시스템 설계

자료처리를 위해 한국정보화진흥원(NIA)에서 제공하는 공공 개방형 클라우드 플랫폼인 PaaS-TA (Platform as a Service TA)를 활용하였다[10]. 병렬 처리를 위해 6대의 서버를 2그룹으로 나눠 클러스터를 구성하고 PaaS 영역을 위해 사설 IP대역 네트워크와 서비스 영역의 공인 IP대역으로 나눠 구성하였다. 서버와 스토리지의 기본 사양은 다음과 같다. 네트워크 구성은 공공기관 네트워크 분리 정책에 따라 외부 인터넷망(external) 1개와 서비스 배포 및 관리를 위해 내부 인터넷망(internal) 대역 2개로 나누었으며 전용구간과 서비스 구간으로 구분하였다. 또한, 보안 강화를 위해 Cloud Foundry를 통해서 배포되는 어플리케이션은 모두 DNS를 통해 접근하도록 구성하였다.

4-2 어플리케이션 구성

APEC기후센터에서 사용하는 PaaS-TA의 서비스 API는 Cloud controller와 서비스 브로커 사이에서 Catalog, provision, un-provision, update provision, bind, unbind 등 통신규약을 구성하였다. 브로커는 http endpoint URI 형식의 Restful로 구현되어 하나 이상의 서비스가 하나의 브로커에 의해 제공될 수 있고, 로드 밸런싱이 가능한 수평적 확장이 제공될 수 있도록 구성되어 있다. 또한, 여러 실행환경의 서비스 인스턴스는 다른 URI 접두사 및 자격증명을 사용하여 하나의 브로커에 의해 지원될 수 있고, 서비스는 서비스 broker API라고 불리는 cloud controller 클라이언트 API를 구현하여 실행환경에서 사용되도록 하였다[11].

표 2. 클라우드 컴퓨팅 시스템 구성

Table 2. Properties for cloud computing system

Type	Properties
OS	vSphere ESXi 6.5.0
Cluster	HP ProLiant DL380 Gen10 Intel XeonGold 5115 CPU
Memory	1,534GB
Storage	123.4TB
Network	SAN Switch 2EA

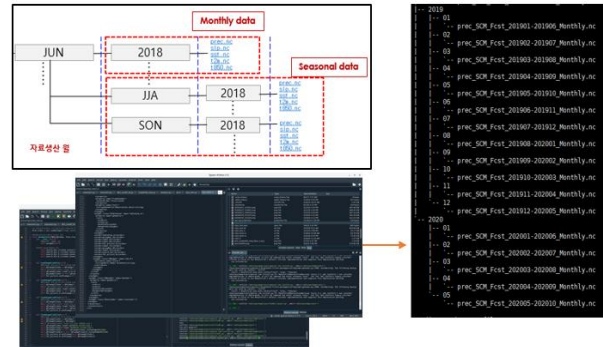


그림 2. CF 규약 기반 계절 예측자료 무결성 검증점검 결과
 Fig. 2. Maintaining the integrity of seasonal forecast data through a program to check compliance with CF Conversion

4-3 자료처리용 메타데이터

계절예측 자료는 연구목적으로 구축된 HPC (High performance computing)에서 생산되고 공유되어 일반 사용자는 자료의 접근이 불가능하다. 그리하여, 계산 영역과 서비스 영역으로 나누어 정기적으로 생산되는 다중모델 앙상블 자료를 개별 모델자료와 같은 위치의 서비스 영역 서버로 동기화시켰다. 동기화 과정에서 자료의 품질점검을 통해 이질적인 속성정보 표준화하였다. 표준화 방법은 OGC에서 제공하는 CF(Climat and Forecast) 규약에 따른 메타데이터 검증 방법을 적용하였다[12]. 자료처리 서비스의 효율성을 높이기 위해 데이터베이스 시스템을 사용하지 않고 오직 구조화된 파일 목록을 사용하였다. 자료검색 및 추출은 구조화된 파일 구성을 이용하여 동기화 작업이 가능하도록 구성하였다. HPC에서 생산된 자료를 서비스용 스토리지로 동기화시키면서 오류파일을 점검하고 전송된 파일의 무결성 검증이 완료되면 미리 정해진 규칙성을 이용하여 파일을 관리하고 누락된 파일은 재 복사하여 최적의 파일 구조 형태를 유지하도록 구성하였다.

V. 서비스 구현

5-1 사용자 웹 인터페이스

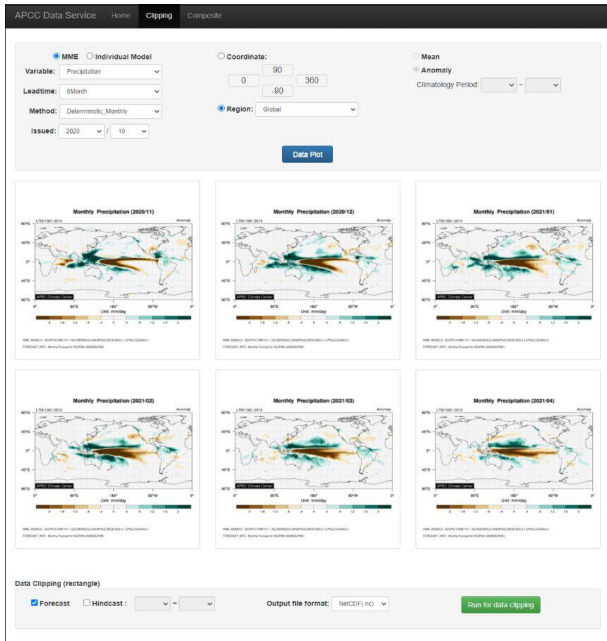


그림 3. 웹 기반 사용자 인터페이스와 시각화 결과
 Fig. 3. User interface that displays the results by lead-time of seasonal forecast data

계절 예측자료 가공 및 추출을 위해 W3C에서 제공하는 웹 표준 프로토콜을 이용하여 사용자 인터페이스를 구성하였다. (https://clips2.apcc21.org/clips) 사용자 인터페이스는 비동기 통신방식을 적용하여 접근의 제약을 해소하고 확장성을 고려했다. 계절 예측자료 가공 및 추출을 위한 화면은 자료검색 및 조건 입력화면, 자료처리 결과 화면 그리고, 자료제공 화면으로 구성하였다. 자료검색 및 조건입력화면에서는 웹사이트에 접속하는 시간을 기준으로 자료의 종류와 최신자료의 목록 파일 스토리지에서 검색하여 표출해 준다. 서비스를 자주 사용하는 사용자들에게 검색조건 입력하는 수고를 덜어주기 위해서다. 사용자가 특정 기간이나 지역을 가공하자 할 경우, 사용자는 원하는 지역, 변수, 기간, 모델, 앙상블 기법 등을 이용하여 자료를 추출할 수 있으며, 개별모델을 이용하여 평균값이나 사용자가 Hindcast 기간을 자유롭게 조절하면서 Anomaly 계산한 결과를 얻을 수 있도록 구성하였다. 주로 사용하는 지역을 미리 정의하여 좌표를 일일이 입력해야 하는 번거로움을 줄이고 정의되지 않은 지역을 선택하고자 할 경우, 좌표입력 화면에서 좌표를 수정하여 지역을 추출할 수 있다. 가공된 자료를 NetCDF나 엑셀 파일형태로 내려받을 수 있도록 구축하였다.

5-2 자료 검색 및 분석영역 지정

자료 분석을 위해서는 사용자가 원하는 대상 자료를 우선 선택해야 한다. 서비스를 위해 사용된 자료는 3개월에서 7개월까지 예측 기간이 다양한 개별 모델 자료와 3개월과 6개월로 통일된 다중모델 앙상블 자료로 나뉜다.

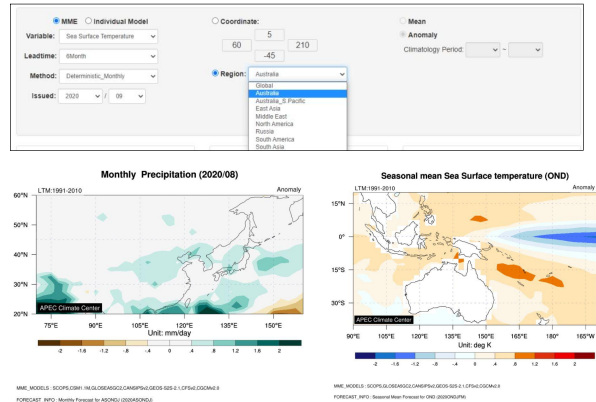


그림 4. 강수와 해수면 온도에 대해 동아시아와 호주지역을 추출한 결과
 Fig. 4. Processing results of East Asian monthly mean precipitation data and sea surface temperature in Australia.

또한, 각 모델을 생산하는 기법에 따라 결정론적 모델 자료와 확률론적 모델 자료로 세분화 된다. 더불어, 각 모델은 편년값을 통해 기후값이나 편차를 선택할 수 있다. 사용자의 편리성을 위해 선택메뉴는 상위에서 하위로 분류시켰다. 생산기관이나 자료의 종류, 변수 종류, 예측 기간, 예측자료생산 기법, 예측 날짜, 추출 좌표정보, 추출 지역 정보, 자료가공 결과를 위한 검색 및 조건입력창이 제공하도록 개발하였다. 멀티 모델 앙상블 기법으로 생산된 계절 예측자료와 참여기관에서 제공하는 개별모델을 바탕으로 검색조건을 바꿀도록 검색 화면을 구성하였다. 사용자는 선택된 자료 종류에 따라 변경되는 하위 메뉴를 선택하여 강수, 해수면 온도, 2m 기온, 해면기압 등 다양한 변수와 6개월, 3개월 다중모델 앙상블, 결정론적 방법이나 확률론적 방법을 선택하고 추출하고자 하는 지역 좌표를 직접 입력하거나 지역을 선택하여 가공처리된 정보를 볼 수 있다. 사용자의 편리성을 위해 지역은 전 지구, 동아시아, 호주, 호주 및 남태평양지역, 중동지역, 북아메리카, 러시아, 남아메리카, 남아시아로 구분하여 제공하지만, 사용자가 특정 지역을 원하는 때는 직접 좌표입력을 통해 자료와 변수를 추출하여 사용할 수 있도록 하였다.

5-3 가공자료 및 추출자료 시각화

기후자료처리 시스템에서 제공되는 변수는 강수, 고도별 기온, 해수면 온도, 해면기압 등으로 처리된 결과를 제공할 때, 각 변수의 특성을 반영한 결과 지도를 제공해야 한다. 범례의 컬러 테이블은 강수와 기온, 편차장과 평균에 따라 다르게 사용되어야 하며, 다중모델 앙상블을 생산하는 기법에 따라 결정론적 방법과 확률론적 방법의 결과 표출 기법을 다르게 구성해야 한다. 결정론적 방법은 이웃하는 값의 윤곽선을 생성(contour)하지만, 확률론적 방법은 좌표정보를 그대로 표시해야 한다.

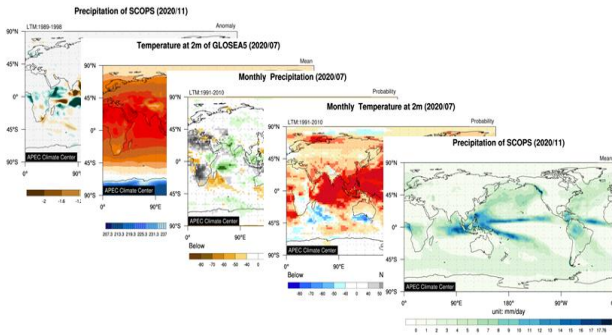


그림 5. 다양한 변수에 대해 가공된 시각화 결과
Fig. 5. Visualization results processed for various variables

전 지구를 기반이 아닌 지역을 추출할 경우, 좁은 영역에서는 값의 간격의 크면 하나의 값을 표출될 수 있어, 최솟값과 최댓값을 기준으로 적절한 간격을 유지하도록 시각화하였다. 강수의 경우, 평균값 정보 제공을 위해 0부터 최댓값을 기준으로 제공하였다.

다중모델앙상블 기법은 기간이나 변수에 따라 조합되는 모델의 수가 변하여 되므로 결과물에 적용된 개별모델 정보를 제공하고 편차를 계산하기 위해 사용된 hindcast 기간을 같이 제공하여 지도기반 시각화 결과에 대한 자세한 정보를 제공하고 있다.

5-4 사용자 맞춤형 편차 계산

전 세계의 모든 기관은 WMO의 권고에 따라 기후 예측을 위해서는 정해진 평년 기간을 사용하고 있다. 평년 기간은 연도에 따라 변경되나, 시스템 개발에 활용된 자료의 평년 기간은 1991년부터 2020년까지 30년의 기간을 적용한 예측 자료이다. WMO에서 권고하는 이유는 각 국가별 평년 기간을 다르게 할 경우, 같은 관측 결과에 대해 특정 지역에서 이상기후 현상으로 분석되는 오류를 방지하기 위해서다. 하지만, 평년 기간을 변경하면서 분석할 경우, 특정 연도에 특이 패턴을 찾는 데 유리할 수 있다. 그리하여 연구자들은 다양한 평년 기간을 적용하여 자신의 지역에서 특이 변화를 찾고자 한다. 이런 분석은 계산자원이 풍부한 대용량 계산 서버나 시스템을 이용해야 하기에 전문연구자가 아니면 쉽게 접근하기 힘들다.

정기적으로 제공되는 APEC기후센터의 다중모델 앙상블 기법을 적용한 계절 예측자료는 편차를 계산하기 위해 정해진 기후 평년값(1991~2020)을 사용하고 있다. 이를 개선하고 사용자들에게 폭넓은 분석 기회를 제공하기 위해 맞춤형 서비스를 위해 사용자들에게 평년값 기간(Climatology Period)을 자유롭게 조절할 수 있도록 서비스를 제공하여 다양한 방법으로 편차를 계산하여 기간에 따른 해당 연도의 예측정보 분석할 수 있도록 알고리즘을 적용하였다.

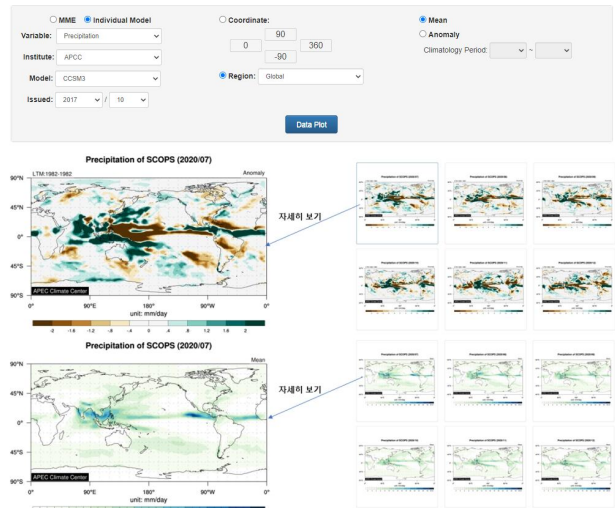


그림 6. 계절 예측자료의 계산된 편차(위)와 평균(아래) 결과
Fig. 6. Anomaly result (above) and Mean result (below) of seasonal forecast data

계산 알고리즘은 변수별로 최적화시켰고 분석기법은 각 모델이 가지고 있는 여러 변수의 특성을 반영하여 평균값이나 평년값에 대해 계산된 편차 정보(Anomaly) 결과를 월별(Lead time)로 표현하며 각 Lead Time을 선택하면 상세정보와 함께 계산 결과를 확인할 수 있다.

APEC 기후센터의 SCOPS 모델의 2020년도 6월에 예측한 강수 자료에 대해 1980년부터 2010년의 Hindcast를 이용하여 전 지구의 강수의 편차(Anomaly)를 계산한 결과와 Mean 평균값 계산 결과로 첫 번째 예측 월인 2020년 7월을 선택하여 자세한 화면으로 확인한 예시이다.

5-5 파일 포맷 변환처리 제공

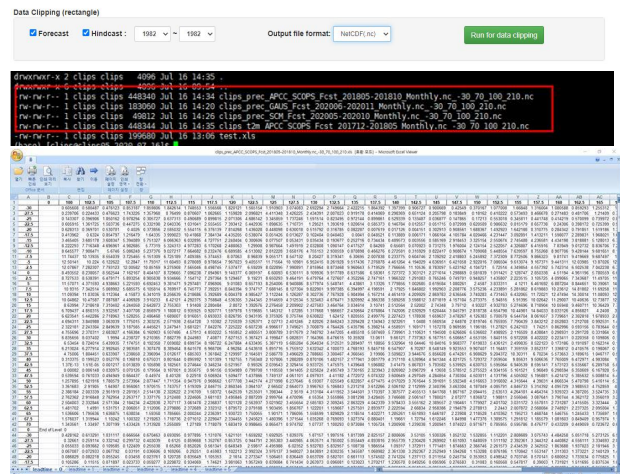


그림 7. 추출된 자료로부터 변환된 파일 포맷 결과
Fig. 7. Data file format conversion of Data extraction and processing results

기후연구 분야에서 가장 많이 사용되고 있는 파일 포맷은 NetCDF 형식으로 시·공간 정보와 많은 변수들이 압축되어 있는 형태로 구성된다. 이런 NetCDF 파일을 사용하기 위해서는 전용프로그램을 설치하고 직접 코딩을 통해 정보를 추출해야 한다. 이런 어려움으로 재난 재해를 대비하는 연구자나 정책결정자는 쉽게 사용하지 못하고 있다. 그리하여, 사용자들이 연구하고자 하는 선택한 변수와 지역 그리고, 계산된 편차 정보 및 평균값을 재활용할 수 있도록 파일형태로 내려받기 서비스를 개발하였다. NetCDF 파일 포맷에 익숙하지 않은 사용자들을 위해 이미지 정보나 엑셀파일 형태로 변환하여 제공하고 공간정보와 시간정보를 하나의 파일에 저장하기 위해 시간정보는 sheet로 구분하고 공간정보는 위도와 경도를 기준으로 테이블 형태로 구성하였다.

VI. 결 론

아시아 태평양 지역은 강수량이 많고 기온이 높은 지역으로 태풍이나 가뭄의 영향을 많이 받는 지역이지만, 이상기후에 대비하기 위한 기후정보 활용이 낮은 지역이다. 그 원인은 대용량 기후자료를 활용할 인프라의 부족과 활용기술의 부족으로 지목된다. 그리하여 고성능 전산 자원이 부족한 국가나 지역의 계절 예측자료 활용을 위해 사용자 맞춤형 계절 예측자료 가공 및 추출 서비스를 개발하였다.

이 연구의 목적은 일반 사용자들이 특별한 계산자원이나 분석기술 없이도 기후자료를 쉽게 분석하고 추출하는 서비스를 개발하는 것이 목표이다. 이를 위해 다음과 같이 구성하였다.

첫째, 시스템 개발을 위해 전 세계의 대표적인 모델 생산기관의 신뢰성 있는 개별 모델예측 자료를 수집하였다. 수집된 자료를 이용하여 APEC 기후센터의 다중모델 앙상블 기법(MME)을 적용하여 다중모델 앙상블 계절 예측자료를 생산하여 제공되어야 할 자료를 확보하였다.

둘째, 표준화된 웹 통신기술을 활용하여 웹 브라우저만 사용하여 모든 서비스가 가능하도록 W3C에서 제공하는 표준화된 웹 표준 규약 코드를 이용하여 웹 표준화 프로그램을 개발하여 경량화 된 웹서비스를 개발하였다.

셋째, 사용자가 직접 명령어를 입력하지 않아도 되는 그래픽사용자인터페이스(GUI) 방식으로 서비스를 제공하여 컴퓨팅 활용 능력이 높지 않은 일반 연구자나 재난 재해를 대비하는 정책결정자를 위한 서비스를 구축하였다.

넷째, 모든 자료처리 및 분석은 APEC 기후센터에서 보유한 고성능 계산자원과 클라우드 컴퓨팅을 활용하여 분산병렬 처리형태로 자료를 처리하고 서비스를 제공할 수 있도록 하였다.

다섯째, 사용자 서비스의 주요 기능은 기후자료 포출 시각화, 지역 추출, 파일 제공 기능과 이해하기 쉬운 다양한 기후자료 제공 파일 포맷 변환기능을 제공하였다.

마지막으로 인증기능을 포함한 기후자료 전송 및 기후자료 서비스를 위한 자료 무결성 검증 기능, 그리고 자료연동 기능을 제공하고 계절 예측자료의 사용자가 원하는 지역, 변수, 기간, 모델, 앙상블 기법 검색 서비스를 위한 웹 인터페이스 개발 및 개별 모델자료의 평균(Mean), 편차(Anomaly)계산 기능을 제공하였다. 사용자가 정의한 주요 지역 목록을 통해 간편한 자료추출 서비스와 사용자가 원하는 특정 좌표를 입력하여 자료추출 기능 제공하였다.

개발된 계절 예측자료 가공 및 추출 서비스를 통해 클라우드 컴퓨팅 기술을 활용하여 분산·병렬처리 기법을 적용하여 대용량 자료의 처리 속도를 향상시키고 사용자가 원하는 처리된 결과물만 추출하여 제공하게 함으로써 자료 분석을 위한 새로운 전산시스템 구축 부담을 덜어 주게 되었다. 또한, 개별 모델 및 다중모델 앙상블 계절 예측자료의 특성상 해당 자료의 형태(NetCDF)에 대해 다소 경험이 부족한 사용자가 있을 수 있고, 사용자가 원하는 지역 및 기후 평년 기간에 대해 제한적으로 제공하고 있는 한계를 극복할 수 있어 사용자가 원하는 자료를 원하는 조건으로 손쉽게 가공 및 추출할 수 있어 폭넓은 연구를 제공할 수 있게 되었다.

이로써, 각 모델에 같은 가중치를 부여하여 종합하는 방식(결정론적 예측)과 개별 모델의 예측확률을 모델별로 가중치를 부여하여 통합하는 방식(확률론적 예측) 등 다양한 처리기술 개발과 서비스 개발의 접목을 통해 개발도상국뿐만 아니라, 기후자료가 필요한 다양한 국가들에 도움이 될 것으로 판단된다.

감사의 글

본 연구는 2020년도 APEC 기후센터의 지원으로 이루어진 결과입니다.

참고문헌

- [1] Jeongmin Han, "Implementation of ESGF Data Node for International Distribution of CORDEX-East Asia Regional Climate Data. International JOURNAL OF CONTENTS, Vol. 17, No. 1, pp. 61-70, March 2021. <https://doi.org/10.5392/IJoC.2021.17.1.061>
- [2] National Centers for Environmental information. National Climate Data Center(NCDC), [Internet]. Available: <http://ncdc.noaa.gov>.
- [3] European Centre for Medium-Range Weather Forecasts. Europe's eyes on Earth(Copernicus), [Internet]. Available: <http://ecmwf.int>.
- [4] APEC Climate Center. Climate Information Processing system(CLIPS) [Internet]. Available: <http://clips.apcc21.org>.
- [5] Kawasaki, A, et al. Data Integration and Analysis System

(DIAS) as a platform for Data and Model Integration: Cases in the Field of Water Resources Management and Disaster Risk Reduction. Data Science Journal, Vol. 17: 29, pp. 1-14. ,September 2018. <https://doi.org/10.5334/dsj-2018-029>.

[6] APEC Climate Center. Open Web Processing Service (OpenWPS) [Internet]. Available: <http://openwps.apcc21.org>.

[7] Yang, Y. B., Song, B. G., Lim, A. Y. & Lim, C. M, Operation and Improvement of the APCC Climate Prediction System, APEC Climate Center, 2019. Technical Report

[8] Krishnamurti, T. N., and Coauthors, Improved Weather and Seasonal Climate Forecasts from Multimodel Superensemble. Science, 285, 1548-1550, 1999. <https://doi.org/10.1126/science.285.5433.1548>

[9] Dandi, A. R., P. A. Pillai, and J. S. Chowdary, Inter-annual variability and skill of tropical rainfall and SST in APCC seasonal forecast models. Clim. Dyn., 56, 439-456, 2020. <https://doi.org/10.1007/s00382-020-05487-w>.

[10] National Information Society Agency. Platform as a Service (PaaS-TA) [Internet]. Available: <http://paas-ta.kr>.

[11] Cloud foundry. Cloud foundry bosh [Internet]. Available: <http://bosh.io/docs>.

[12] Cho Jaepil, Jung imgook, Cho Wonil, Lee Eun-Jeong, Kang Daein and Lee Junhyuk, Suggestion of User-Centered Climate Service Framework and Development of User Interface Platform for Climate Change Adaptaion. Journal of climate change Research 2018, Vol.9, No.1, pp,01-12, 2018 <http://doi.org/10.15531/KSCCR.2018.9.1.01>.



한정민(Jeong-Min Han)

2003년 : 공주대학교 대학원 (교육정보학석사-교육정보학)
2008년 : 충북대학교 대학원 (공학박사-컴퓨터공학)

2012년~현 재: APEC기후센터 기후분석과 선임연구원

※관심분야 : 기후자료처리(Climate Data Processing), 기후분석시스템(Climate Analysis System), 인공지능(AI) 등



임창묵(Chang-Mook Lim)

2011년 : 국립 강원원주대학교 대기환경과학 (이학사)
2014년 : 국립 강원원주대학교 대학원 (이학석사-대기환경과학)

2015년~현 재: APEC기후센터 예측운영과 연구원

※관심분야 : 계절예측(Seasonal Prediction), 기후자료처리(Climate data processing) 등