

TRAVEL REPORT FORM

출장보고서

결 재	선임연구 원	팀장	부서장	원장
	12/21	12/21	12/21	12/24
협 조	한정민	민영미	유진호	권원태
	연구원			
	12/21			
	최재원			

I. Travel Overview 출장개요

1. Traveler(s) 출장자

Department 소속	Position 직위(직급)	Name 성명	Note 비고
예측운영팀	선임연구원	한정민	
	연구원	최재원	

2. Travel Period 출장기간

- 2018년 12월 2일 ~9일

3. Occasion and destination 행사 및 출장지

- 행사명: 8th Annual ESGF2018 F2F Meeting
- 출장지: Marriott Marquis, Washington DC, USA

*ESGF: Earth System Grid Federation *F2F : Face to Face

II. Major Activities 주요업무 수행내용

1. Main Contents and Activities 주요내용 및 활동

가. 주요내용

- ESGF F2F Meeting은 ESGF (Data, Index, IDP, Compute) 노드와 관련된 시스템 개발자 및 운영자들이 직접 만나서 열리는 회의로 각 기관에서 운영하고 있는 ESGF 운영 상태, 이행계획, 컴퓨팅 서비스, 아키텍처 등 ESGF 관련 전반적인 기술 공유하는 자리로 ESGF 분야 선진기술 습득, LLNL, NSC, IPSL, DKRZ, JPL 등 전문가와 의견교류를 통해 APCC에서 구축한 ESGF in KOREA의 데이터 노드 운영에 선진기술을 반영하고자 함.



그림 1 발표회장 사진

나. 논문발표

- “How can the Coordinated Regional Climate Downscaling Experiment(CORDEX) community facilitate dissemination and analysis of the next generation climate simulations in the Cloud?”를 주제로 포스터 발표하였음(공동저자).

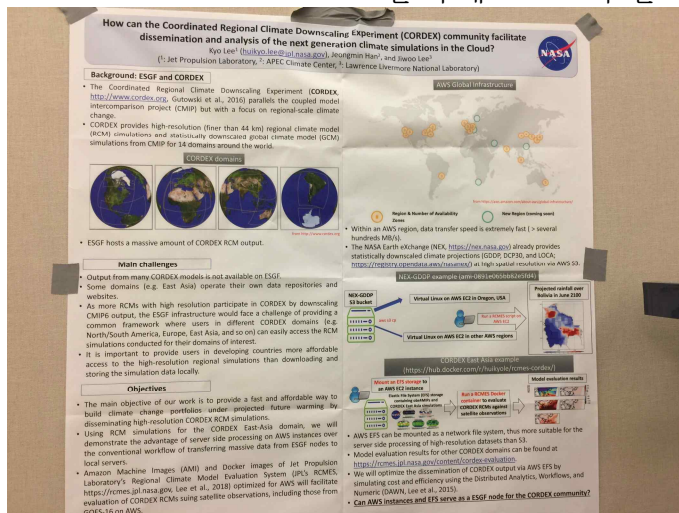


그림 2 논문발표 포스터 사진

다. 데이터 노드 연동 업무

- 개요: APCC에서 구축 중인 ESGF in KOREA 데이터 노드는 CORDEX-EA 자료를 배포하기 위한 데이터 전송 노드로 CORDEX자료를 담당하는 Tier1 그룹의 스웨덴 노드 서버와 연동되어야 함. 시스템 연동의 문제를 갖고 있어 스웨덴 노드 담당자를 직접 만나 시스템 구축의 제약사항과 문제를 해결하였음.
- 내용: APCC의 데이터 노드와 스웨덴 노드의 연동 실패의 원인(ESGF 공식배포 버전(2.7.1)으로 설치되었으나 스웨덴 노드는 업그레이드 버전(2.8 RC1)으로 설치되어 있었음)을 확인하고 Keypairing 과정에 에러가 발생하였던 원인을 찾아 수정하고 APCC의 서버에서 만들어진 cert key를 기반으로 NSC(스웨덴) 노드 담당자(Mr Prashanth)와 키 등록 및 생성을 통해 새로 발급된 hostkey를 이용하여 key 매핑작업을 진행하였음. 2019년도에 업그레이드 버전이 배포될 예정이며 자료를 배포하기 위한 작업을 컨퍼런스가 끝난 이후에 NSC스웨덴 노드 담당자와 함께 공동 작업하였음.

- 1) National Aeronautics and Space Administration
- 2) Lawrence Livermore National Laboratory
- 3) infrastructure for the european network for earth system modelling

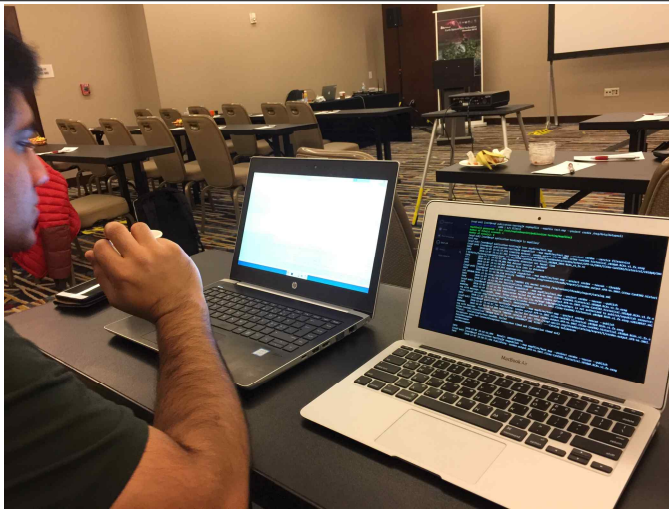


그림 3 스웨덴 노드 전문가와 공동 작업하는 사진

라. 공동 연구제안 관련 논의

- 개요: JPL(이희교 박사)와 APCC가 공동으로 NASA를 대상으로 공동 프로젝트 제안의 가능성에 대해 논의하였으며 공동 연구 주제 및 관련 내용을 협의 하였음.
- 내용: 현재 ESGF는 대용량 기후 데이터를 공동으로 관리하고 교환하는 역할을 수행하고 있지만 사용자들이 대용량의 데이터를 활용하여 분석하기에는 전산자원(네트워크, 저장 공간)의 제약이 따름. 이런 문제를 해결하는 방안으로 저장공간에 구매받지 않고 자료전송 속도를 높이는 방안을 제시하고자 하였음. 지역기후 모델인 CORDEX-EA 자료의 경우, 수요자가 동아시아 국가이며 동아시아 국가에서 대부분 활용한다는 점과 CORDEX 도메인 14개와 AWS(Amazon Web Service)의 Cloud 지역 센터의 서비스 영역이 일치되는 점을 착안하여 CORDEX-EA 자료를 대상으로 AWS서울 센터를 활용하여 동아시아 지역 기후모델 평가시스템 (RCMES : Rgional Climate Model Evaluation System) 구축을 NASA에 제안하기로 하였음. 프로젝트 공고와 제안 시기는 19년 1월에서 2월 사이로 예상됨.

마. 발표청취 및 기술동향 파악

- ESGF F2F Meeting에서는 ESGF 컴퓨팅 서비스 분야, 컴퓨팅서비스와 머신러닝, ESGF에서의 CMIP6 서비스, ESGF 서비스 연동, ESGF 분석어플리케이션, Rethinking of ESGF 아키텍처로 구성되었으며 관련 전문가들이 모여 발표와 논의를 하였음.
- NASA¹⁾에서는 ESGF의 컴퓨팅 서비스 분야에서는 Web Processing Service 기술을 적용하여 NASA에서 개발한 Earth Data Analytics Service에 대한 머신러닝 처리 기술에 대해 발표를 하였으며 LLNL²⁾에서는 Django와 pyWPS를 기반으로 개발한 CWT의 아키텍처에 대해 소개, 컴퓨팅 노드를 이용하기 위한 인증, 보안, 자원 배분에 대한 내용을 설명하였음. IS-enes³⁾에서는 자신들이 개발한 CDS(Climate Data Store)의 CMIP5, CORDEX 자료를 처리하기 위한 컴퓨팅 서비스의 요구조건과 구성에 대해 발표하였음.

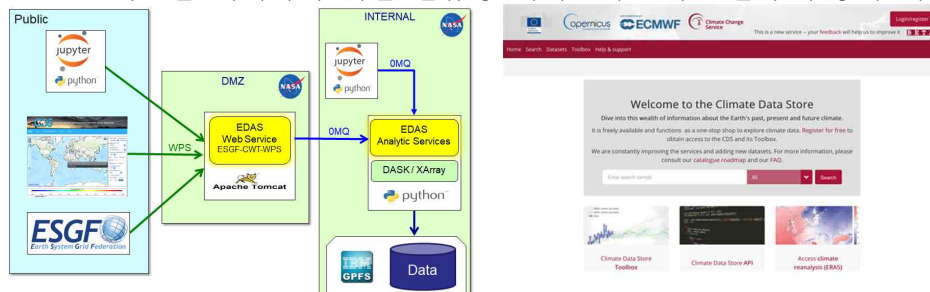


그림 4 EDAS 아키텍처(좌)와 CDS 인터페이스(우)

ESGF 데이터를 이용한 머신러닝 분야에서는 CMCC⁴⁾에서 개발한 Ophidia의 고성능 데이터 분석 및 머신러닝 기술에 대해 소개하였고 LLNL에서는 ConvLSTM을 이용하여 극한

기후현상과 허리케인 진로 예측에 대해 머신러닝 기법을 적용한 결과를 발표하였다. CRIM⁵⁾에서는 Earth Observation Exploitation Platform에서 머신러닝 기법을 적용하기 위한 testbed를 개발하고 개발 과정에 대한 내용을 소개하였음. PSL⁶⁾에서는 CMIP6 제공을 위한 Errata서비스와 SYNDA서비스, CDNOT(CMIP Data Node Operations Team)를 소개하고 앞으로의 워크플로를 발표했음. CMCC에서는 데이터 사용에 대한 통계 시스템을 구축하여 Dashboard를 이용하여 다양한 사용정보를 통계화하는 툴을 소개하였음. 또한, DKRZ⁷⁾ 데이터 citation 서비스에 대해 소개하였음.

ESGF 서비스 연동 분야에서는 OGC⁸⁾의 WMS(Web Map Service), WMTS(Web Map Tile Service), WFS(Web Feature Service), WCS(Web Coverage Service), WPS(Web Processing Service) 등 많은 서비스와 KML, GML, GeoPackage, GeoTIFF, NetCDF, HDF 등 다양한 기후자료 포맷에 대한 OGC의 기술의 ESGF 활용에 대해 소개하였음.

2. Relevance to APEC Climate Center's Activities 결론 및 소감

- 18년도 APCC의 ESGF 데이터 노드를 구축하면서 “대용량 기후데이터에 대한 국제적 교류와 활용”이라는 측면에서 궁금했었던 기후자료 처리 컴퓨팅 기술에 대한 고민을 많이 해소할 수 있었던 자리였음.
- ESGF는 각 국가를 대표할 수 있는 기후 관련 대표기관들의 참여로 만들어진 Global 서비스로 기후데이터(CMIP, CORDEX)에 대한 IT기술의 총집합체라고 보여지며 APCC의 기후데이터 기반 서비스(ADSS, CLIPs, OpenWPS)에 대한 통합서비스 구축을 위해 ESGF 기술과 아키텍처를 설계 단계에서부터 활용하고자 함.
- 그림5는 CMCC에서 작성한 ESGF 데이터 활용에 대한 지역별 통계로 아시아 지역에서 ESGF데이터를 가장 많이 사용하고 있음을 알 수 있음. 그리하여 APCC의 ESGF의 연합체 참여는 아시아 지역을 대표할 수 있는 데이터 서비스 기관으로서의 위상을 높일 수 있는 계기가 되었다고 생각됨.
- 또한, JPL과 공동으로 NASA에 제안할 “동아시아 지역 기후모델 평가시스템 구축” 사업이 이뤄진다면 ESGF 데이터 노드 운영과 지역 기후모델 평가 시스템 구축을 통해 아태지역 연구자들에게 다양한 서비스를 제공할 수 있을 것으로 판단됨.



그림 5 CMCC에서 작성한 지역별 ESGF데이터 사용 통계

3. Suggestions and Remarks 건의사항

- APCC의 ESGF(Earth System Grid Federation) 연합체 참여는 기관의 입지를 확고히 할 수 있는 계기가 될 수 있어 ESGF시스템에 대한 전문기술을 습득하고 확장 서비스 개발 등 관련 업무에 적극적으로 참여하고자 함.

III. References (Presented and Collected Materials) 주요 수집자료

(with attachment of any information or report in case of attendance of conferences, workshops and meetings) 학술대회, 워크숍, 회의 등 참석 시 관련 정보 및 문서 첨부

- ESGF 관련 전문가 연락처
- ESGF 관련 기술정보 및 발표자료 수집

-
- 5) Centre de recherche informatique de montreal
 - 6) Institut Pierre Simon Laplace
 - 7) German Climate Computing Center
 - 8) Open Geospatial Consortium

- Look for conference AGENDA and ABSTRACT updates on the [ESGF Website](#).
- [BOOK LODGING NOW!!!](#)
- The registration fee is waived for remote participants. PLEASE JOIN US VIA [WEBEX](#) "FREE"- Access code: 807 447 408 - Meeting password: esgf

REGISTER and BOOK LODGING BY 10/03/2017



The Earth System Grid Federation (ESGF) Conference is an international collaboration for the software that powers most global climate change research, notably assessments by the International Panel on Climate Change. The conference is to continue distributed data federation development for the climate community at large that is supported by DOE, Office of Science.

ESGF 2018 will be held at the Marriott-Marquis, Washington, DC. Listed below is an overview of the scheduled events:

December 3, 2018 (2:00p - 4:00p) - Pre-conference Registration

(5:00p - 6:00p) - Social - Meet & Greet (NO HOST)

December 4 – 7, 2018 (8:00a-5:00p) - Conference Convened

December 7, 2018 (8:30a-10:00a) - ESGF Executive Committee Breakout Meeting

(10:00a - 12:00p)- ESGF Develop. Team Report Back on Conference Findings

(1:30p - 5:00p) - Optional; General Code Sprint

Where.....
Marriott - Marquis
901 Massachusetts Ave NW
Washington, District of Columbia 20001
USA
+1(202) 824-920

Summary	Information Agenda
Abstracts - DUE by September 30, 2018	First Section • Look for conference AGENDA and updates on the <u>ESGF Website</u>.
Registration Costs - Due by October 10, 2018	Monday December 3rd: Arrival
Cancellation Policy	Tuesday December 4th
Hotel and Washington, DC Area Information	Morning • General Aspects • ESGF Status (review of 2016 implementation plan, review action items from ESGF F2F 2017, status of ESGF working teams) • Keynotes from ESGF board and funding agencies
Agenda	Afternoon • ESGF Compute Services
Social Activities - December 3, 2018	Late Afternoon • Poster session
Location	Wednesday December 5th
Contact Us	Morning • CMIP6 Services in ESGF
	Afternoon • ESGF in Analysis Applications
	Thursday December 6th
	Morning • Rethinking of ESGF Architecture
	Afternoon • Topics from ESGF Working Teams
	Friday December 7th
	Morning (parallel sessions) • Working Teams Side Meetings • ESGF-XC F2F



October 4, 2018

Jeongmin Han
APEC Climate Center (APCC)
Busan, South Korea

Dear Jeongmin,

Lawrence Livermore National Laboratory will host the Earth System Grid Federation (ESGF) 2018 Conference (ESFG 2018 F2F) on December 3 – 7, 2018 at the Marriott Marquis, 901 Massachusetts Ave NW, Washington, DC 20001. It is our pleasure to invite you to attend and to give a possible featured address.

The Earth System Grid Federation (ESGF) 2018 Conference (ESFG 2018 F2F) is an international collaboration for the software that powers most global climate change research, notably assessments by the International Panel on Climate Change. The conference is to continue distributed data federation development for the climate community at large that is supported by DOE, Office of Science. (ESGF, www.esgf.llnl.gov)

The discussions to take place during the conference will be important to our goals for developing software components, which will build a stronger federated infrastructure. The workshop discussions will also cover present day standards needed for future collaborations between the sites.

Please use the ESGF Conference page at <https://www.cvent.com/events/8th-annual-esgf-f2f-conference-esgf-2018-/registration-d2c9d372ed56433cab3e5b80c9541c50.aspx?fqp=true> to find links to forms, hotels, and travel arrangements. Please complete the registration and book your room as soon as possible as space at the conference and hotel is limited. The deadline for registration is October 12th.

Please send participation forms and questions to: Alexander (Sasha) Ames – ames4@llnl.gov and Angela Jefferson – jefferson9@llnl.gov or you may call +1 (925) 422-5932 for immediate assistance.

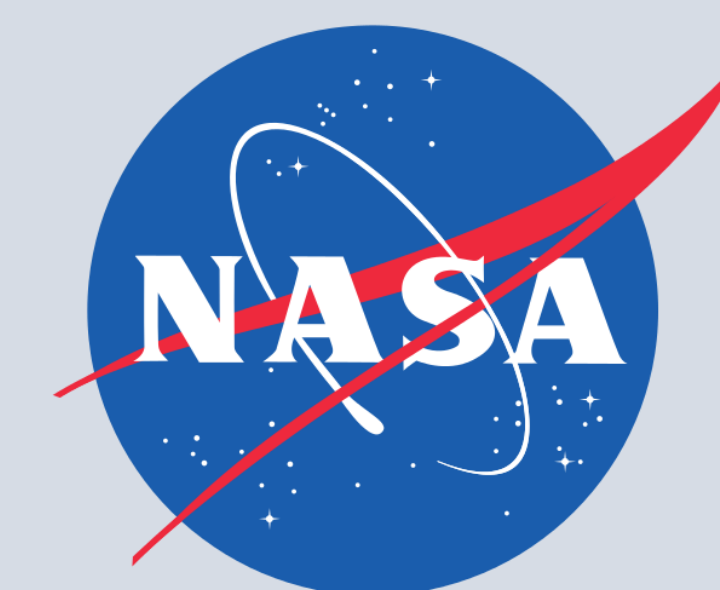
As more information becomes available, particularly in terms of your agenda and presentation scheduling, we will let you know. We look forward to seeing you at the ESFG 2018 F2F.

Regards,

A handwritten signature in black ink, appearing to read 'Sasha Ames', with a long horizontal flourish extending to the right.

Alexander (Sasha) Ames

How can the Coordinated Regional Climate Downscaling Experiment (CORDEX) community facilitate dissemination and analysis of the next generation climate simulations in the Cloud?



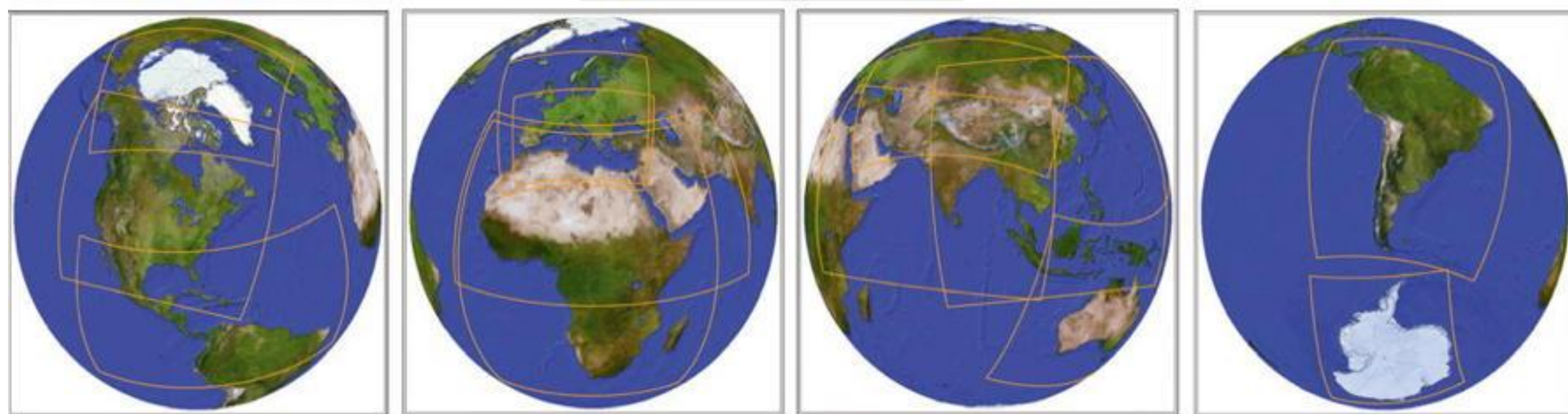
Kyo Lee¹ (huikyo.lee@jpl.nasa.gov), Jeongmin Han², and Jiwoo Lee³

(¹: Jet Propulsion Laboratory, ²: APEC Climate Center, ³: Lawrence Livermore National Laboratory)

Background: ESGF and CORDEX

- The Coordinated Regional Climate Downscaling Experiment (**CORDEX**, <http://www.cordex.org>, Gutowski et al., 2016) parallels the coupled model intercomparison project (CMIP) but with a focus on regional-scale climate change.
- CORDEX provides high-resolution (finer than 44 km) regional climate model (RCM) simulations and statistically downscaled global climate model (GCM) simulations from CMIP for 14 domains around the world.

CORDEX domains



from <http://www.cordex.org>

- ESGF hosts a massive amount of CORDEX RCM output.

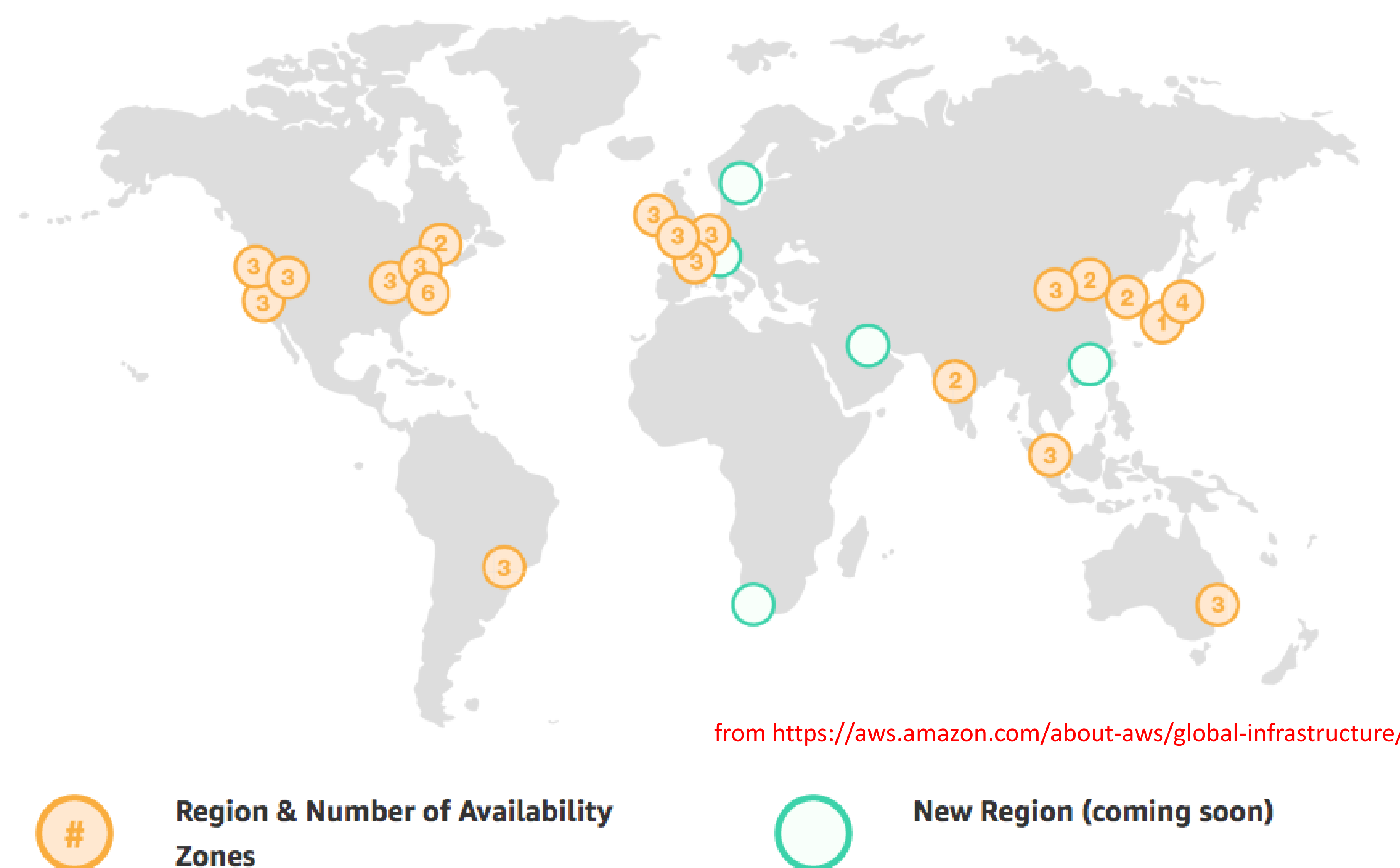
Main challenges

- Output from many CORDEX models is not available on ESGF.
- Some domains (e.g. East Asia) operate their own data repositories and websites.
- As more RCMs with high resolution participate in CORDEX by downscaling CMIP6 output, the ESGF infrastructure would face a challenge of providing a common framework where users in different CORDEX domains (e.g. North/South America, Europe, East Asia, and so on) can easily access the RCM simulations conducted for their domains of interest.
- It is important to provide users in developing countries more affordable access to the high-resolution regional simulations than downloading and storing the simulation data locally.

Objectives

- The main objective of our work is to provide a fast and affordable way to build climate change portfolios under projected future warming by disseminating high-resolution CORDEX RCM simulations.
- Using RCM simulations for the CORDEX East-Asia domain, we will demonstrate the advantage of server side processing on AWS instances over the conventional workflow of transferring massive data from ESGF nodes to local servers.
- Amazon Machine Images (AMI) and Docker images of Jet Propulsion Laboratory's Regional Climate Model Evaluation System (JPL's RCMES, <https://rcmes.jpl.nasa.gov>, Lee et al., 2018) optimized for AWS will facilitate evaluation of CORDEX RCMs using satellite observations, including those from GOES-16 on AWS.

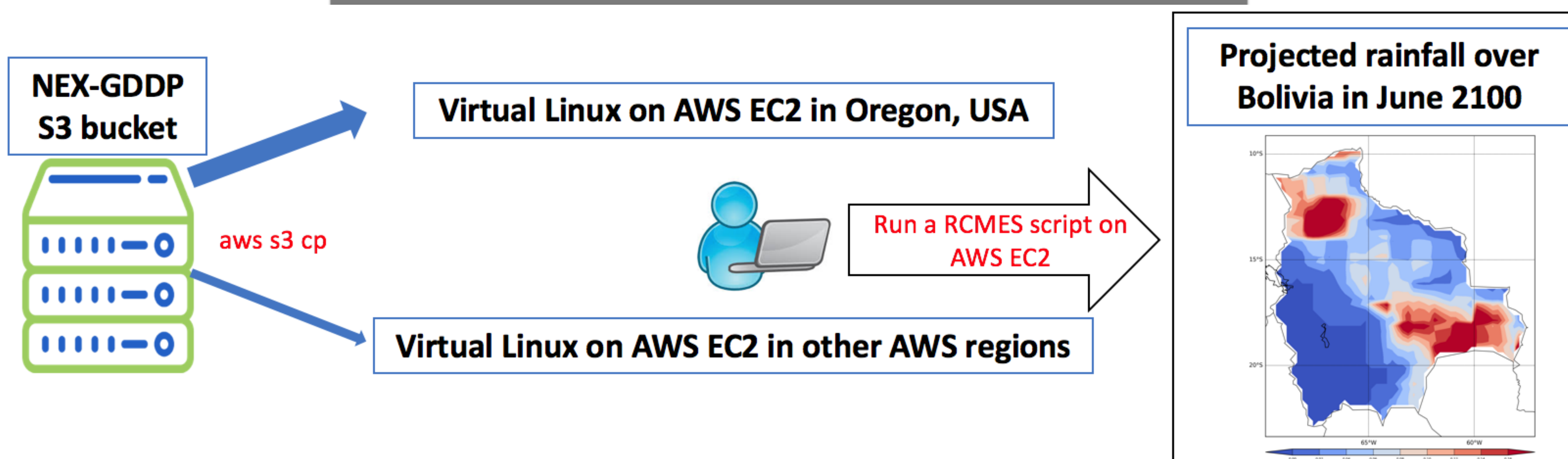
AWS Global Infrastructure



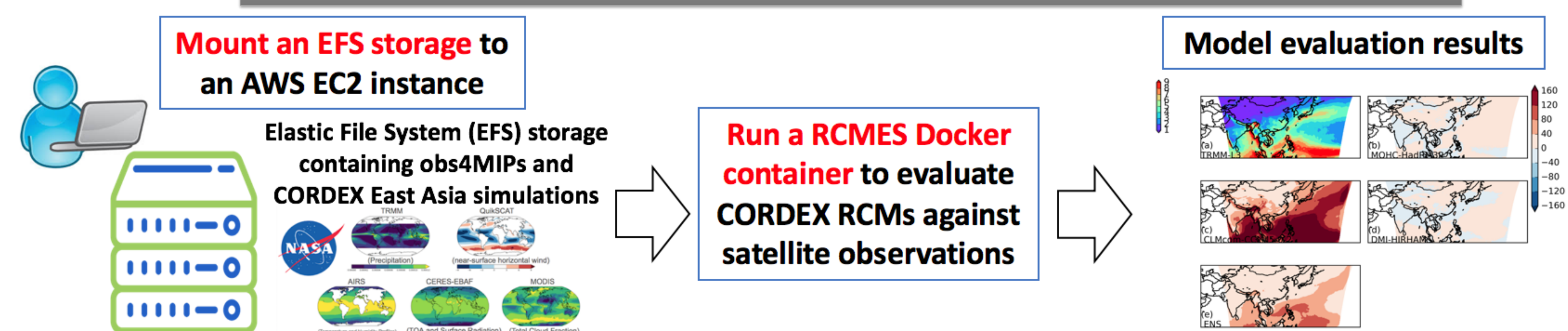
from <https://aws.amazon.com/about-aws/global-infrastructure/>

- Within an AWS region, data transfer speed is extremely fast (> several hundreds MB/s).
- The NASA Earth eXchange (NEX, <https://nex.nasa.gov>) already provides statistically downscaled climate projections (GDDP, DCP30, and LOCA; <https://registry.opendata.aws/nasanex/>) at high spatial resolution via AWS S3.

NEX-GDDP example (ami-0891e065bb82e5fd4)



CORDEX East Asia example (<https://hub.docker.com/r/huikyole/rcmes-cordex/>)



- AWS EFS can be mounted as a network file system, thus more suitable for the server side processing of high-resolution datasets than S3.
- Model evaluation results for other CORDEX domains can be found at <https://rcmes.jpl.nasa.gov/content/cordex-evaluation>.
- We will optimize the dissemination of CORDEX output via AWS EFS by simulating cost and efficiency using the Distributed Analytics, Workflows, and Numeric (DAWN, Lee et al., 2015).
- Can AWS instances and EFS serve as a ESGF node for the CORDEX community?**

References

Gutowski, W. J., et al. (2016), WCRP Coordinated Regional Downscaling Experiment (CORDEX): a diagnostic MIP for CMIP6, Geosci Model Dev, 9(11), 4087-4095, doi: 10.5194/gmd-9-4087-2016.
Lee, H., L. Cinquini, D. Crichton, and A. Braverman (2015), Optimization of system architecture for Big Data analysis in climate science, paper presented at Big Data (Big Data), 2015 IEEE International Conference on.
Lee, H., A. Goodman, L. McGibbney, D. E. Waliser, J. Kim, P. C. Loikith, P. B. Gibson, and E. C. Massoud (2018), Regional Climate Model Evaluation System powered by Apache Open Climate Workbench v1.3.0: an enabling tool for facilitating regional climate studies, Geosci Model Dev, 11(11), 4435-4449, doi: 10.5194/gmd-11-4435-2018.

© 2018 California Institute of Technology. Government sponsorship acknowledged.

