Atmospheric River Tracking Method Intercomparison Project

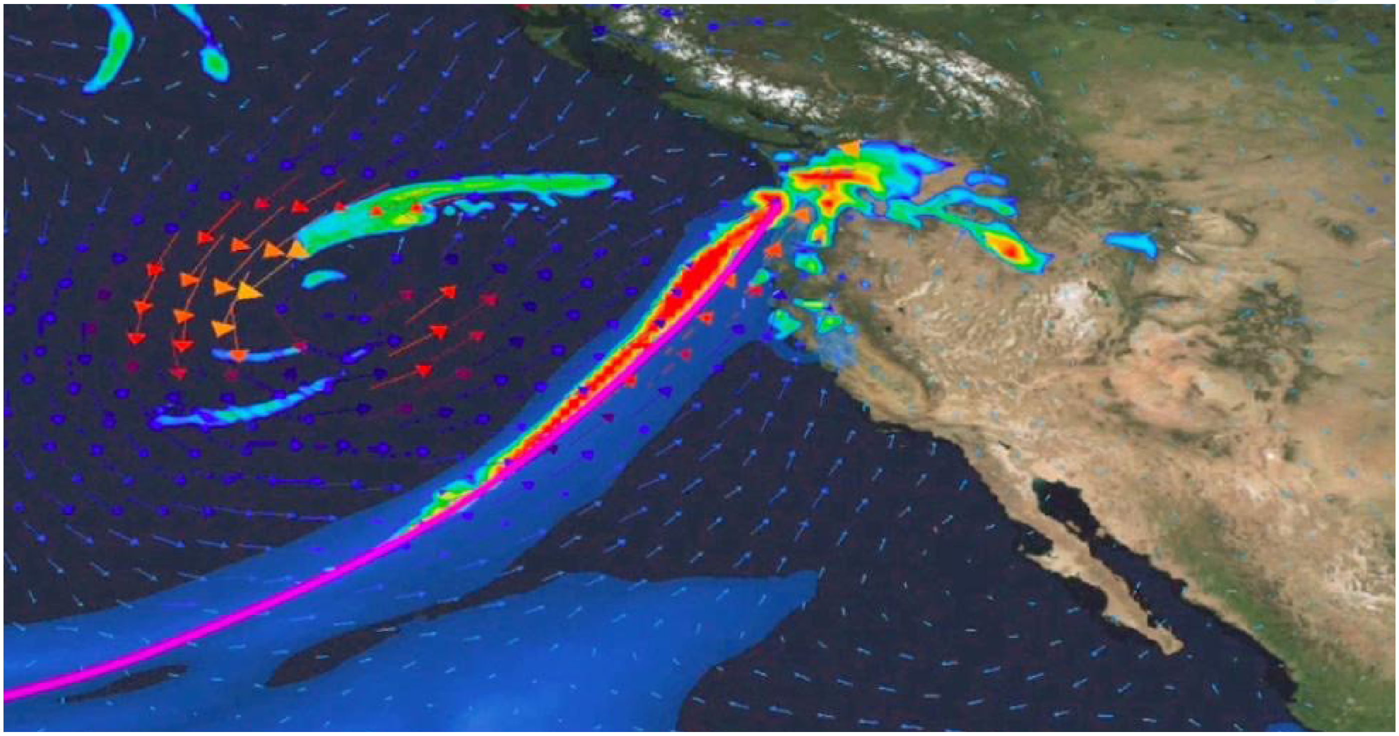
ARTMIP

Report of the Second ARTMIP Workshop

Gaithersburg, Maryland, April 23-24, 2018

DOE/SC-0194

U.S. DEPARTMENT OF ENERGY



Based on a vision to improve the characterization and predictability of atmospheric rivers (ARs) on both weather and climate time scales, the Atmospheric River Tracking Method Intercomparison Project (ARTMIP) aims to quantify the uncertainty in AR climatology (e.g., frequency, duration, and intensity), precipitation, and related impacts that arise because of different AR tracking methods, and uncertainty in how these AR-related metrics may change in the future. The ARTMIP also aims to provide guidance regarding the advantages and disadvantages of different AR tracking methods and which of these methods are best suited to answer certain scientific questions. Finally, the ARTMIP will develop an online repository of data for future use in research.

Supported by

Renu Joseph, Regional and Global Model Analysis

U.S. Department of Energy

Office of Science

Office of Biological and Environment Research

Co-Chairs, Committee, and Writing Team

Christine A. Shields, National Center for Atmospheric Research

Jonathan J. Rutz, National Oceanic and Atmospheric Administration

L. Ruby Leung, Pacific Northwest National Laboratory

F. Martin Ralph, Scripps Institution of Oceanography

Michael Wehner, Lawrence Berkeley National Laboratory

Travis A. O'Brien, Lawrence Berkeley National Laboratory

Roger Pierce, National Oceanic and Atmospheric Administration

About the cover

Main Photo: Image of an atmospheric river frontal boundary taken during the Atmospheric River Reconnaissance 2018 Field Campaign.

Globe: Atmospheric rivers from the ARTMIP catalogues for February 7, 2017.

EXECUTIVE SUMMARY

Atmospheric rivers (ARs) have for the last decade been broadly recognized for their global and regional significance in mediating energy and water cycles. As a distinct large-scale circulation feature, ARs can be characterized by long, narrow bands of atmospheric moisture transport from the tropics to higher latitudes. They are both an important source of water supply to coastal regions as well as a hazard because of their predisposition to increase flood risk. However, predicting ARs poses a significant challenge to the scientific community, as they exhibit large variability from subseasonal-to-interannual time scales. On multi-decadal and longer time scales, it is unclear how the frequency and intensity of ARs will change with a warming climate. Hence ARs play a significant role in understanding predictability of water cycle variability and extremes as well as becoming a factor in assessing risks to future infrastructures.

There is a large body of literature that has explored ARs from both global and regional perspectives and from timescales spanning hourly to centennial. Despite significant progress in understanding ARs, significant gaps remain, e.g., in characterizing the dynamical and thermodynamical processes associated with ARs, improving predictions of ARs at weather and subseasonal-to-seasonal timescales, and understanding and quantifying the impacts of ARs on the design and management of built infrastructures. Furthermore, the consensus definition of ARs remains essentially qualitative, and investigators have had to rely on heuristic definitions of ARs for quantitative studies.

In the First Atmospheric River Tracking Method Intercomparison Project (ARTMIP) workshop held in May 2017, it was recognized that (1) there are many heuristic AR tracking algorithms employed in the literature, (2) our quantitative—and possibly qualitative—understanding of ARs and their impacts may depend on the details of the algorithm used, (3) no data set yet exists to systematically explore the impact of AR tracking algorithm choice on scientific results, and (4) the broader scientific community therefore lacks any formal guidance on the advantages and disadvantages of different AR tracking methods. The First ARTMIP workshop defined and launched a multi-tiered intercomparison experiment designed to fill this community need. The first tier of ARTMIP was aimed at understanding the impact of AR algorithms on quantitative baseline statistics and characteristics of ARs, and the second tier includes sensitivity studies designed around specific science questions, such as reanalysis uncertainty and influences associated with climate change.

The Second ARTMIP workshop was convened on April 23-24, 2018, as a two-day event, and following the completion of Tier 1 of the ARTMIP effort: an experiment in which most existing AR tracking algorithms were run on a standardized data set of reanalyzed weather conditions. The goals of the Second ARTMIP workshop were to provide a forum to:

1. discuss gaps and emerging opportunities for advancing the science and tracking of ARs,
2. discuss analyses of the Tier 1 data set,
3. synthesize the results and implications of the Tier 1 analyses,
4. use this information to define the experimental designs for the various Tier 2 experiments, and
5. work towards developing systematic analyses and evaluation of the advantages and disadvantages of different AR algorithms for various scientific questions.

The two-day Second ARTMIP workshop was sponsored by the U.S. Department of Energy and included 36 participants from various U.S. federal agencies/programs, national and international universities, and DOE national laboratories. In addition to an opening presentation by the US Global Change Research Program (USGCRP) that is tasked to coordinate federal investment, the workshop included detailed agency-level presentations from the Department of Energy, the National Oceanic and Atmospheric Administration, and the National Aeronautics and Space Administration. These were followed by a set of presentations on current activities and science gaps in AR research. The presentations also helped frame the discussion for the following sessions, which were structured to meet the above goals.

Second ARTMIP Workshop Results and Outcomes

Presentations and discussions during the workshop led to the following consensus statements:

- The various AR algorithms that are used by the community are based on different sets of descriptors that characterize ARs, e.g., ranging from the broad outline of the region with high water-vapor transport, to the details of the core region in which the transport is most intense.

- The different empirical methodologies that are based on different descriptors or characterizations of the regions within ARs has led to different ways to statistically characterize ARs, that in turn has led to differences in the predicted intensity, duration, frequency, size, and tropical-to-high-latitude vapor transport.
- AR algorithms can broadly be grouped, or "clustered" by specific characteristics, for example: global versus regional, use of absolute versus relative thresholds, and use of information about temporal continuity.
- Some statistics and characteristics of ARs differ significantly among these groups of algorithms.
- Algorithms that are appropriate for identifying ARs globally and in mid-latitudes are not appropriate for identifying ARs entering high-latitude regions, due to the precipitous drop in atmospheric water vapor transported by ARs as they move poleward.
- Specification and timelines for the following Tier 2 experiments were defined to expand on the Tier 1 experiment:
    - Intercomparison of AR tracking algorithms with multiple reanalysis data sets;
    - Intercomparison of AR tracking algorithms on high-resolution simulations of present and projected future climate;
    - Intercomparision of AR tracking algorithms on output from phase 5 and phase 6 of the Coupled Model Intercomparison Project.

Key Gaps and Research Priorities

Gap: It will be necessary to understand the apparent spread in the statistics and characteristics of ARs among AR algorithms and allow recommendations to be made for which AR algorithms are more appropriate for addressing different scientific questions.

Research Priorities: It will be critically important to define a categorization of AR algorithms according to the regions within individual ARs that the algorithms tend to identify as part of the ARs. AR identification and tracking methods should also be grouped into "clusters" on the basis of key algorithm criteria such as geometric requirements. It must also be clearly communicated that these expert-developed methods were developed to answer different questions: therefore it makes sense that answers vary. This information adds context to ARTMIP results that might otherwise be perceived as having too much spread.

Gap: There is still a major outstanding question of whether our quantitative—and possibly qualitative—understanding of how ARs may change in the future depends on the details of the algorithm used.

Research Priorities: Active multi-institutional participation and collaboration in the Tier 2 experiments analyzing ARs in high-resolution climate simulations and Coupled Model Intercomparison Project Phase 5/6 (CMIP5/6) simulations is key to resolving whether and how our understanding of ARs and their changes in the future may depend on the tracking algorithm, model resolution, and other uncertainty factors.

Gap: AR identification and tracking methods are based on different integrated vapor transport (IVT) thresholds (either as an absolute value or anomaly/percentile) to act as a lower bound for identifying ARs, and AR intensity for each method invariably scales as a function of where this lower bound is set. A major challenge is communicating to stakeholders how the choice of this lower bound, and other criteria, affect the impacts associated with ARs identified by different methods.

Research Priorities: The definition of AR intensity can be informed and refined by ARTMIP, but different methods will likely persist. Ultimately, ARTMIP should focus on developing novel ways of communicating uncertainty information that arises from differences in method criteria to stakeholders and the public.

Gap: Uncertainties in AR impacts, in part defined by ARTMIP analysis, can be communicated to stakeholders such as local governments and water managers.

Research Priorities: Applying ARTMIP data and analysis to the impacts communities such that extreme precipitation and runoff metrics can be given a range of uncertainty given the diverse set of algorithms and metrics available in the literature. The ARTMIP community can provide guidance to stakeholders to help discern which algorithms metrics are most appropriate for them.

Gap: There is a need for systematic analysis and quantification of uncertainties, and the implication of those uncertainties, in AR climatology and AR-related impacts that arise from the use of different AR tracking methods throughout the literature. These analyses, which will inform both the science and application communities, can only be facilitated by the maintenance of an organized data set based on these methods.

Research Priorities: Catalogues should be made available to the scientific community to address science questions that will inevitably arise from ARTMIP analyses. Data management and storage is required for Tier 1 and 2 catalogues and will be housed on the Climate Data Gateway at NCAR as the project progresses.

Gap: Beyond ARTMIP, more research is needed to improve forecasting of the precursors and evolution of ARs for longer lead times, develop metrics and diagnostics to understand and quantify model forecast and simulation skill, improve understanding of the sources of predictability of ARs from subseasonal to multi-decadal timescales, and improve communication geared towards the general public and stakeholders in need of AR forecast and projections.

Research Priorities: ARTMIP provides an important framework for systematic analysis of ARs and their regional and global impacts, focusing on uncertainty associated with AR tracking methods. This should open the door to further advance AR science to bridge the gap between research and societal needs. Improving forecasting and modeling of ARs and their impacts requires more observations, particularly observations from field studies that complement satellite and global reanalyses data. High-resolution data sets are also needed in support of high-resolution modeling that is becoming more feasible with advances in high-performance computing and numerical methods such as regional refinement using unstructured grids. Numerical experiments using a hierarchy of models can be effective for studying AR predictability, AR impacts, and ARs in the paleoclimate records.

Photos illustrating the wide range of water levels at Oroville Dam and associated challenges faced in managing water in an environment of extremes, for which variability is driven by atmospheric rivers.

TABLE OF CONTENTS

Executive Summary i

1. Background 1

1.1 What are atmospheric rivers?. 1

1.2 Current research, science gaps, and programmatic relevance 1

1.3 What is ARTMIP and why is it important? 3

1.4 ARTMIP Tier description 4

2. Science questions, summaries, status, plans 5

2.1 Science questions. 5

2.2 Tier 1 summary 5

2.3 Tier 1 status and future analysis 7

2.4 Tier 2 summary and status. 7

2.5 Future plans and timeline 8

2.6 Gaps and research priorities for ARTMIP 8

3. Beyond ARTMIP: *Emerging needs and opportunities for atmospheric rivers research* 10

3.1 AR forecasting and observations 10

3.2 AR climate research 11

3.3 AR impacts. 11

3.4 Gaps and research priorities beyond ARTMIP 12

4. References 12

5. APPENDICES 14

Appendix A1 14

Appendix A2 16

Appendix A3 16

Appendix A4 17

Appendix A5 19

1. BACKGROUND

1.1 What are atmospheric rivers?

Atmospheric rivers (ARs) are dynamically driven, filamentary structures of atmospheric vapor transport that have been associated with ~90% of poleward water vapor transport outside of the tropics, despite occupying only ~10% of the longitudes. They are often associated with extreme winter storms and heavy precipitation along the western coasts of mid-latitude continents. They have the ability to produce major flooding events but they may also relieve droughts. Because ARs play an important role in the global hydrological cycle as well as regional water resources, understanding how they vary from subseasonal to interannual time scale and how they may change in the future is critical to advancing understanding and prediction of regional precipitation.

A broad definition of an AR (Figure 1) can be found in the AMS Glossary of Meteorology (added May 2017). This was developed through a consensus-driven process described in Ralph et al. (2018a).

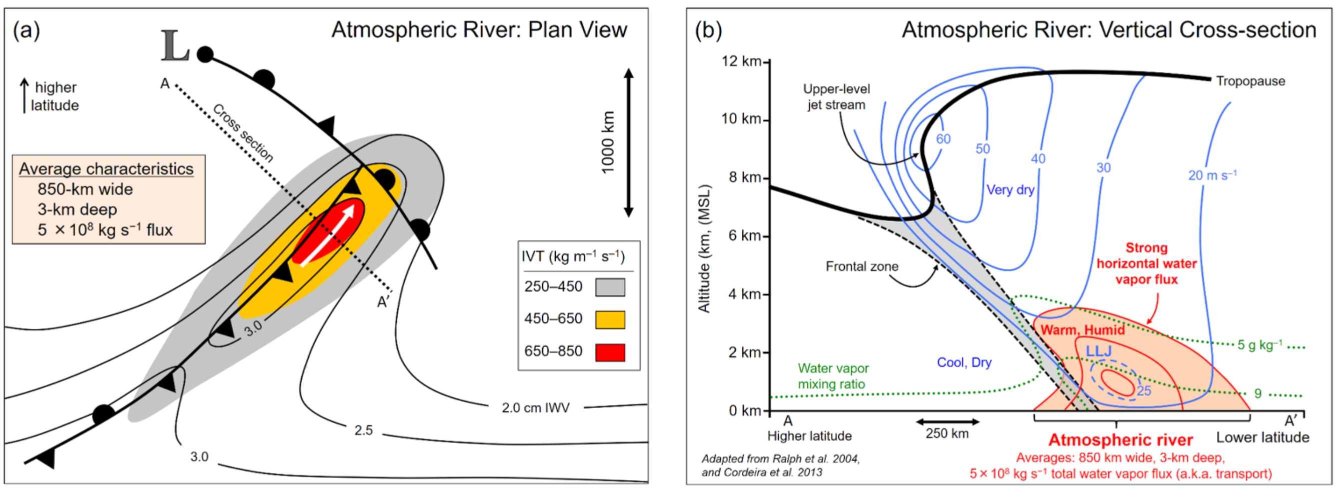


Figure 1. From the AMS Glossary of Meteorology and the result of a consensus-driven effort to define an atmospheric river (a) shows the AR Plan View and (b) shows the vertical cross-section.

1.2 Current research, science gaps, and programmatic relevance

Interest in AR research is growing and robust for communities involved in research on weather and climate across a variety of scales: from synoptic-scale forecasts to centennial-scale climate projections. Recent field programs and missions serve to improve AR forecasts and understanding of the meteorological factors and aerosols and cloud physics that influence AR precipitation, which will ultimately support the impacts communities, local governments, and general public. Academic research into AR processes and mechanisms continues to deepen, including understanding potential changes in ARs for near-term and future climates. AR predictability is also emerging as a key focus spanning the continuum from subseasonal to seasonal, and seasonal to decadal to centennial, timescales.

Because of the scientific challenges in predicting ARs and associated precipitation and hydrologic extremes and the significant societal impacts of ARs, improving understanding and modeling of ARs is an important objective that aligns with interagency efforts such as the water cycle coordination under the US Global Change Research Program's Integrated Water Cycle Group (IWCG). The IWCG, with participation from the Department of Defense (DOD), DOE, Department of the Interior (DOI), Environmental Protection Agency (EPA), National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA), National Science Foundation (NSF), and the Department of Agriculture (USDA), focuses on global change-related research and activities related to the integrated water cycle from a multi-scale perspective and seeks an end-to-end approach spanning fundamental research to application. Coordinated research on ARs may forge increased interactions between the weather and climate communities to advance the study of extreme events from complementary perspectives.

Droughts in the western U.S. have important impacts on U.S. energy systems. As a few ARs may provide enough precipitation to mitigate droughts, improving predictions of ARs may influence energy-relevant decisions. Hence, improving understanding and modeling of ARs and associated impacts supports the Integrated Water Cycle Scientific Grand Challenge of DOE Biological and Environmental Research Climate and Environmental Sciences Division to advance understanding of the integrated water cycle. Addressing this grand challenge will improve the predictability of the water cycle and reduce associated uncertainties in response to short- and long-term perturbations, which is important for developing sustainable solutions to water, energy, and environment. Hence DOE sponsored the two-day Second ARTMIP workshop to provide a forum to discuss gaps and emerging opportunities for advancing the science and tracking of ARs.

The Second ARTMIP workshop included 36 participants from various U.S. federal agencies/programs, national and international universities, and U.S. national laboratories. The workshop began with a series of big-picture presentations spanning the weather/forecast and climate modeling communities, focusing on advances and gaps in AR science. Recent cross-agency collaborations were highlighted, including AR reconnaissance missions that involved flights, forecasting, and data assimilation aimed at improving AR forecasts (Figure 2). These AR reconnaissance missions were able to make use of moist adjoint method to identify target area where better observations can improve forecast of heavy rain on the U.S. west coast. Dropsonde data were successfully integrated into global models to improve AR forecast. Stakeholders, including decision makers of water management, transportation, and agriculture (in addition to the general public), all rely on accurate forecasts and reliable projections for how the statistics and characteristics of ARs may change in the future. There is a critical need for improved AR predictions as well as communication of those predictions.

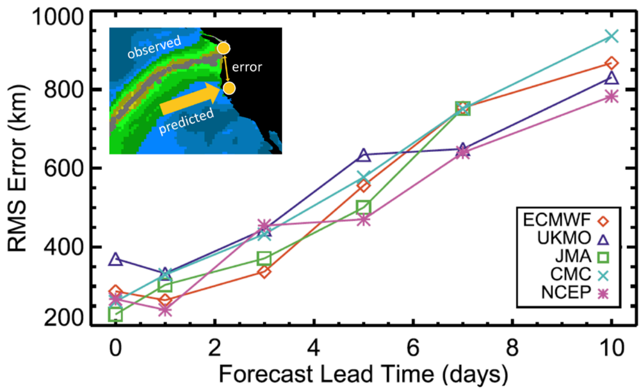


Figure 2. *400-km AR landfall position forecast error at a 3-day lead time. AR forecast skill assessment establishes a performance baseline. (Wick et al. 2013). From "AR Reconnaissance "presentation by F.M. Ralph (Scripps/CW3E).*

At the workshop, a new, proposed set of definitions for scaling ARs was debuted and greeted with enthusiasm. The proposed AR scaling categories ARs based on intensity (maximum integrated vapor transport) and duration to characterize the strength and impacts of ARs (Ralph et al. 2018b). Analogous to rating hurricane winds between the scale of 1 to 5, scaling ARs may improve the communication of ARs and their impacts to the stakeholders and general public. From cat 1 to cat 5, the impacts of ARs vary from "beneficial" to "hazardous."

Advances in modeling ARs in earth system models were also presented and discussed. A frontier in global earth system modeling is increased model resolution with the purpose of more accurate regional simulations. Increasing horizontal resolution improves the simulation of jet position and strength (Lu et al. 2015) (Figure 3), which have been linked to model biases in simulating AR frequency (Gao et al. 2016). Increasing resolution is also important for improving simulation of AR precipitation, snowpack, and runoff in regions of complex terrain. Ensemble modeling has been proven to provide more useful projections of future changes in ARs by better characterizing uncertainty associated with atmospheric internal variability and biases in simulating ARs and extreme precipitation (Hagos et al. 2016). There is a need to further explore the use of global variable-resolution models for balancing the needs for increasing model resolution and increasing the number of ensemble members using regional refinement to simulate ARs and their future changes.

Understanding the sources of predictability of ARs is important for improving predictions of extreme precipitation and flooding associated with ARs. A recent study of variability of extreme precipitation in the western U.S. confirmed the strong relationship between ARs and winter extreme precipitation (Dong et al. 2018). However, the study found that sea-surface temperature (SST) forcing accounted for only about 20% of the variances of winter extreme precipitation. The remaining 80% of extreme precipitation variations was associated with internal atmospheric variability, which is harder to predict than the slowly evolving SST. The internal atmospheric variability is related to circumglobal waveguide in the mid- to high-latitude atmosphere, which can interact with anomalous convection in the tropical western Pacific. Understanding these processes and their predictability will be important for improving prediction of winter extreme precipitation in the western U.S. from subseasonal to interannual timescales.

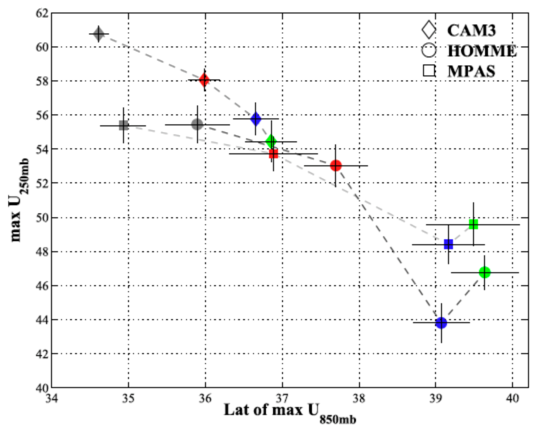


Figure 3. *Latitude of the zonal wind maximum at 850 hPa (in degree) and the 250 hPa maximum zonal wind (in $m\ s^{-1}$) in simulations with different dynamical cores and model resolutions. The subtropical jet location and strength are sensitive to resolution with a hint of convergence at ~50 km grid spacing (Lu et al. 2015).*

Because AR impacts are typically regional, AR studies often focus on specific geographic areas around the globe. There has been extensive work focusing on regions such as western North America, the United Kingdom, and the Iberian Peninsula. Lesser, but equally important, are other regions such as the southeast U.S., western South American coast, and polar regions. Although ARTMIP workshop/meeting participation has been heavily weighted toward western North America, the group included active participation from European and South America. European topics dovetail with those focusing on the western U.S. and include the origin of moisture sources, the relationship of ARs to the dynamics of extra-tropical cyclones, and the predictability of ARs across time scales including future climate projections (Gimeno et al. 2016). Detecting ARs over the poles pose different challenges than ARs in the mid-latitudes and require retooling algorithms typically developed for North America or Europe. Thresholds often associated with AR tracking in mid-latitudes do not sufficiently capture moisture transport into the poles. Arctic and Antarctic ARs potentially feed back to the global energy and water cycles by influencing glacial mass balance with either extreme melting or accumulation events. AR impacts in high latitudes are an important area of interest as suggested by recent publications describing anomalous snow accumulation over East Antarctica by an AR (Gorodetskaya et al. 2014) and the impacts of ARs on Arctic sea ice (Yang and Magnusdottir 2017).

1.3 What is ARTMIP and why is it important?

The goal of ARTMIP is to understand and quantify uncertainties in AR science based on choice of detection/tracking methodology. The climatological characteristics of ARs, such as AR frequency, duration, intensity, and seasonality, all strongly depend on the method used to identify ARs. It is, however, the precipitation attributable to ARs that is perhaps most strongly affected, and this has significant implications for our understanding of how ARs contribute to regional hydroclimate now and in the future. Understanding the uncertainties and how the choice of detection algorithm impacts quantities such as precipitation (Figure 4) is imperative for stakeholders such as water managers, city and transportation planners, agriculture, or any industry that depends on global and regional water cycles information for the near term and into the future. Understanding and quantifying AR algorithm uncertainty is also important for developing metrics and diagnostics for evaluating model fidelity in simulating ARs and their impacts. Resolving this uncertainty is also critically important for addressing grand challenges in understanding the *Integrated Water Cycle, including extreme events, and High-Latitudes processes* (U.S. DOE, CESD 2018).

Scientific focus on ARs has grown exponentially over the last decade, and consequently many methods have been developed to track ARs. Quantifying the uncertainties associated with AR tracking methods and understanding the implications of their differences is the motivation for ARTMIP. The chart below (Figure 5) summarizes the many different algorithmic approaches found in current literature by broadly categorizing the variety of parameters used for identification and tracking, and then listing different types of choices available per category.

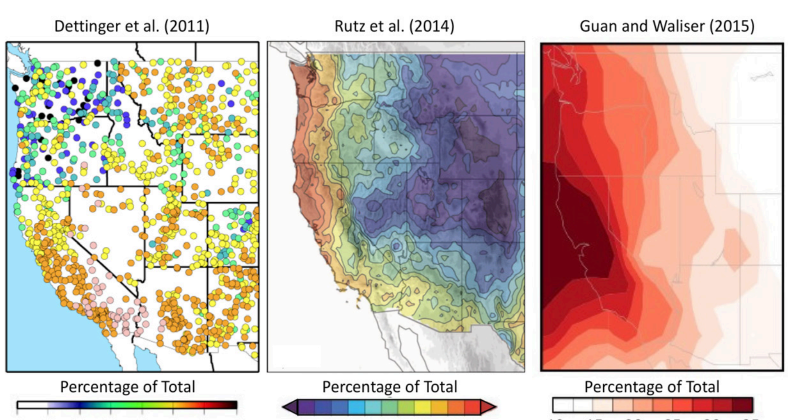


Figure 4. *Shown (left and center) is the fraction of cool-season precipitation attributable to ARs from Dettinger et al. (2011) and Rutz et al. (2014). (Right) As in left and center, but for annual precipitation. These studies use different AR identification methods, as well as different atmospheric reanalyses and observed precipitation data sets.*

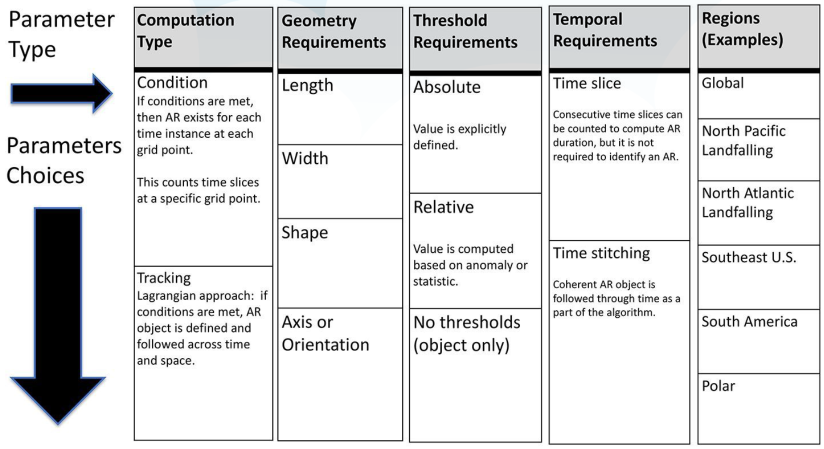


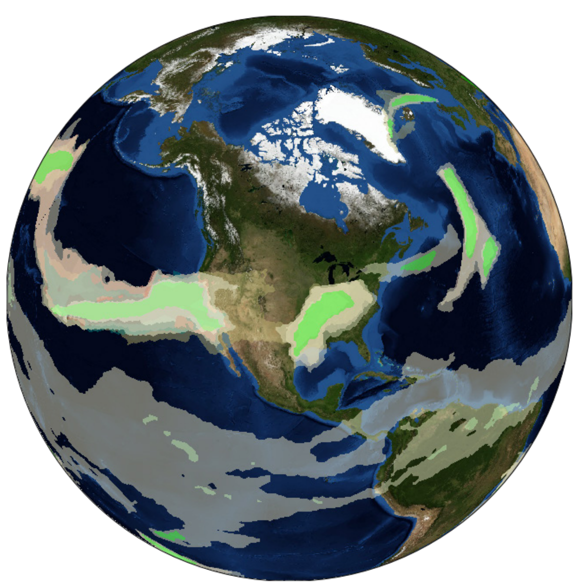
Figure 5. *Schematic diagram illustrating the diverse components of AR detection algorithms. Parameter types commonly applied include computation type, geometric moisture threshold, and temporal requirements, as well as algorithms developed for global ARs or specific geometric regions. (Shields et al. 2018.)*

ARTMIP provides the framework necessary to compare each method in a consistent and objective way. It is divided into two phases, Tier 1 and Tier 2. The phased approach loosely groups catalogues (the data set created by each method) by the science questions of interest. Catalogues are created by the participating groups that run their algorithms on a common data set. Full details can be found in Shields et al. (2018), which explains the project's experimental design and goals. This paper also highlights preliminary results from a successful 1-month proof-of-concept test that was conducted by ARTMIP participants to test the veracity of our design. In this proof-of-concept trial run, the month of February, 2017 was chosen because it was a very active AR month for the U.S. west coast.

1.4 ARTMIP Tier description

Tier 1 is the first phase of ARTMIP. All catalogues are created using the MERRA-2 Reanalysis from January 1980 to June 2017. Contributing to the Tier 1 catalogues is required for participation in ARTMIP and will provide the baseline for all subsequent catalogues generated by Tier 2, the second phase of ARTMIP. Over 12 algorithms have been used to detect and track ARs in Tier 1. Each algorithm produces output that includes a timestamp, latitude, longitude, and a binary tag for AR identification, i.e., 0 for no AR present at the given pixel and 1 for AR present at the given pixel. Tier 1 catalogues have been submitted and we are beginning the process of analyzing the data. The first day of the workshop included presentations and discussions around Tier 1 results.

Tier 2 catalogues mimic the design of Tier 1, but catalogues will be grouped into different science topics. Topics that have been outlined thus far include: climate change using high-resolution model output, climate change based on CMIP5/6 data, and sensitivity to different reanalysis products. The second day of the workshop was dedicated to establishing the details and timeline for Tier 2 catalogues and topics.



2. SCIENCE QUESTIONS, SUMMARIES, STATUS, PLANS

2.1 Science questions

Understanding the uncertainties, and the implications for those uncertainties, in AR science based on algorithmic choice for AR-identification is the driving question guiding ARTMIP. Many science questions follow from this basic goal and have been discussed amongst the ARTMIP participants in depth. A sample of these questions follows:

- How do metrics such as frequency, duration, intensity, and precipitation associated with ARs change from one algorithm to the next?
- Which algorithms are best suited for addressing AR impacts?
- Are there major differences between global versus regional tracking?
- Can AR tracking methods be equally useful for forecasts versus climate projections?
- How do algorithmic choices impact the representation of AR dynamics?
- How and why do different algorithm choices change our understanding of ARs now and into the future?
- Do global models represent AR characteristics and processes accurately, and how do AR tracking methods influence this assessment?
- What are the drivers for AR genesis based on ARs tracked using different methods?
- What forecast variables and forecast skill are most useful for stakeholders?
- Do AR tracking methods affect assessment of forecast skill and hence communication of the usefulness of AR forecasts to stakeholders?

2.2 Tier 1 summary

Previously, many scientists ran their own AR identification and tracking methods over different regions, using different data sets, and during different periods of record. ARTMIP Tier 1 involved each participant running his or her respective method on global data, from the MERRA v2 reanalysis, from January 1980 through June 2017 (Gelaro et al. 2017). The resulting data set is unique in terms of scope and intuitive, being composed of 1s (positive AR identification) and 0s (negative AR identification) at each grid point and time step. This allows for comparing and contrasting and quantifying uncertainty in results between methods based on a common region, a common data set, and a common period of record.

An overview of ARTMIP's Tier 1 analysis will be described in a full-length, peer-reviewed journal article and will be led by ARTMIP co-chair, Jonathan Rutz. The Tier 1 overview paper will focus on a few key metrics and aims to quantify the uncertainty in the climatological characteristics of ARs (e.g., frequency, duration, intensity, seasonality, and relationship to precipitation) that arise from the use of different AR identification and tracking methods. At the Second ARTMIP workshop, substantial discussion also took place on the preliminary results and are summarized below.

AR frequency and seasonality

Given that the diverse algorithms produce a range of frequencies, visualizing and comparing these methodologies can be done most effectively by grouping, or clustering, the algorithms into different categories. "Clustering" (e.g., by absolute thresholds, relative thresholds, length requirements, etc.; see Figure 5) reduces the amount of data shown and facilitates interpretation. The spread in AR frequency can be explained, in part, by the different algorithmic choices, and by identifying what part of the AR, spatially and temporally, the algorithms are targeting.

Another common metric is to identify the month of maximum frequency for each algorithm (Figure 6). In general, there is remarkable agreement in the climatology for the month during which the highest number of ARs make landfall among the methods for both western coast of North America (West) and Europe (East).

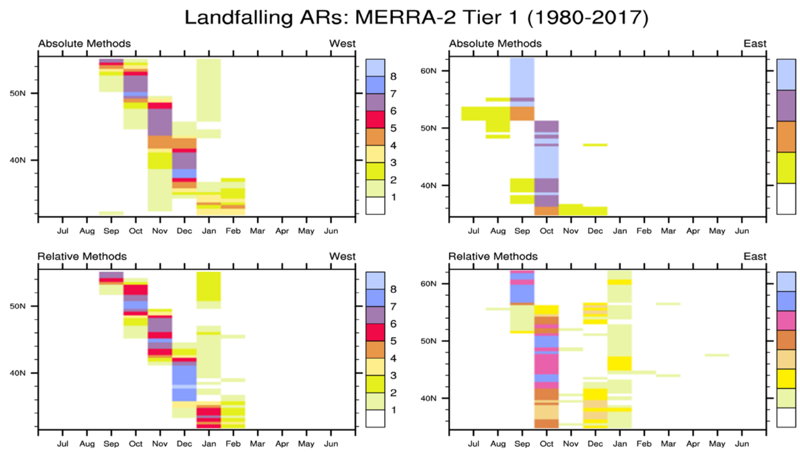


Figure 6. *Agreement amongst methodologies for the month of maximum frequency for landfalling ARs. The west coast of North America (West) and Europe (East) are shown. Plotted in color pixels are the number of methods using absolute (relative) moisture thresholds for a given month. The color bar maximum value corresponds to the number of methods possible for this category. For example, for absolute methods applied to Europe (East) (top right panel), the total possible number of methodologies is 4 and is plotted in light blue. All 4 absolute/east methods agree in September as the month with maximum frequency of landfalling ARs for latitudes 55-62N. MERRA-2 for 1980-2017 were used.*

AR duration

Comparing AR duration amongst the different algorithms can be accomplished by examining the overlap between AR detection methods for AR cases making landfall on global west coasts. Examples of this can be found in 1) the experimental design ARTMIP paper (Shields et al. 2018), which focused on analysis for one month (February 2017), and 2) Figure 7 applied to the full Tier 1 data (courtesy of Ashley Payne, Univ. of Michigan). This type of analysis nicely shows a time series of the methods that do and do not identify ARs either at a point or along a certain coastline, allowing for quick interpretation of consistency between algorithms. Alternatively, computing duration climatologies for predefined time intervals allows for comparison of clustered categories.

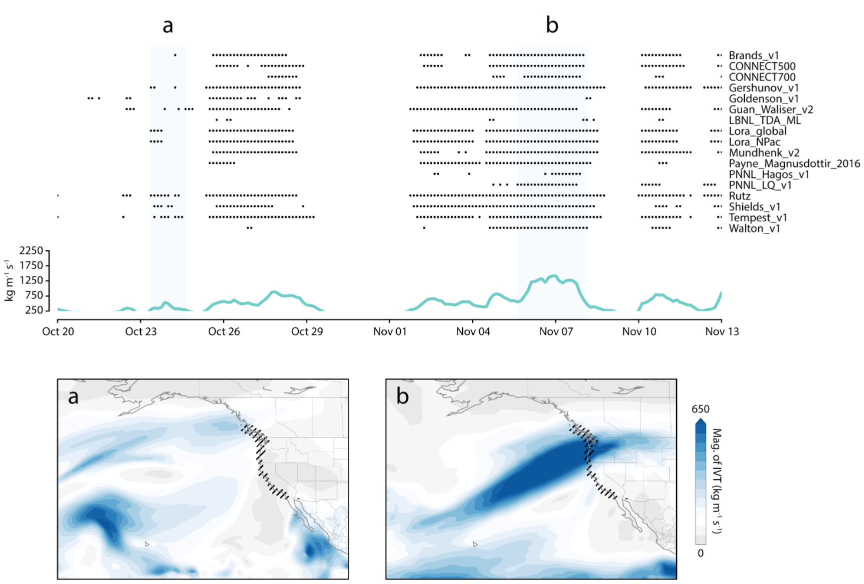


Figure 7. *Time series of the magnitude of IVT along the U.S. west coast in teal for October 20 – November 13th 2006. The dots at the top of the figure show time steps with AR conditions along the coastline for each of the algorithms. Two events are highlighted in light blue to show the range of agreement between the algorithms. Panel (a) shows a composite in IVT of a relatively smaller event that was identified in only eight algorithms (averaged between Oct 23 09 UTC and Oct 24 15 UTC) and panel (b) shows a much stronger event captured by nearly all algorithms (averaged between Nov 05 15 UTC and Nov 08 03 UTC). The region used to determine landfall is hatched in both (a) and (b). The first event falls along the northern edge of the landfalling domain and highlights that differences in the spatial extend of the AR id might influence landfall rates.*

AR intensity

Diagnosing and comparing AR intensity can be tricky when comparing the different algorithms. In some regards, the intensity of the ARs described by a particular method will be inherent by the definition of the algorithm itself, specifically for those that use a moisture threshold. For example, a method that requires a limit of 500 kg/m/s IVT will inherently represent more "intense" ARs than those that allow for lower thresholds, such as 250 kg/m/s. Nonetheless, AR intensity is an important metric to evaluate. We will need to not only compare algorithms but define what we mean by intensity. An initial attempt at defining "intensity" was presented at the workshop. This analysis calculates how closely each AR identification or tracking method comes to replicating the canonical results of Zhu and Newell (1998), such that ~90% of poleward water vapor transport is accomplished in ARs comprising ~10% of global circumference at any given latitude. This method of definition has inspired ARTMIP participants to pursue related analyses along the lines of "AR efficiency" (see discussion of Tier 1 overview paper, section 2.4).

Analysis comparing a subset of detection algorithms to a specific observational point has been done along the southern California coast (Bodega Bay) (Ralph et al. 2018c). Using analysis techniques similar to the Ralph et al. 2018c study, ARTMIP catalogues can also be similarly validated.

2.3 Tier 1 status and future analysis

Diagnosing AR frequency, duration, and seasonality (i.e., month of maximum frequency) is straightforward. Results based on each method can be shown on cross-sections along selected transects to highlight changes with latitude. These transects can be located along coastlines (e.g., the North American and European west coasts) to depict landfalling ARs, or over the oceans (e.g., North Pacific and Atlantic Oceans) to depict patterns away from topographic influence. Additionally, results can be plotted globally or regionally on plan-view maps, denoting median or mean, and some measure of spread.

Assessment of AR intensity, and the relationship between ARs and precipitation, are more nuanced. Comparing AR intensity across methods is challenging because so many methods are intrinsically based on the magnitude of some variable exceeding a given threshold (e.g., IWV > 20 mm, IVT > 250 kg m^{-1} s^{-1}). However, ARTMIP participants are currently exploring novel ways to discuss AR efficiency, or the ratio of water vapor transport of identified ARs to the spatial footprint of identified ARs (obviously, the strictest methods will produce the highest ratios, and setting a "floor" on AR size will need to be considered).

The relationship between ARs and precipitation will be assessed in two ways. The first is by considering the fraction of certain rare (e.g., top-decile) precipitation events attributable to ARs, which reveals the role that ARs play in high-impact weather events. The second is by considering the fraction of total precipitation attributable to ARs, highlighting the role that ARs play in long-term hydroclimate. Both rely on a method for attributing precipitation to ARs, but many such methods already exist (e.g., Dettinger et al. 2011; Rutz et al. 2014; Guan and Waliser 2015), making selection or alteration of a method straightforward. Both fractions listed above can be plotted globally or regionally on plan-view maps, denoting median or mean, and some measure of spread. This study will use MERRA v2 precipitation data, despite known issues, for the benefit of self-consistency with the rest of the analysis.

AR case studies will examine differences in how the various methods identify ARs in time and space, and the implications of these differences for the attribution of precipitation to ARs. Ideally, both a strong AR event and a weak AR event will be examined, highlighting sensitivity of method-based uncertainty as a function of AR strength. In addition, these case studies will lend themselves to comparison of AR duration as a function of AR identification method at selected locations experiencing these AR events.

Finally, many of these analyses will proceed by clustering methods based on their thresholding criteria (e.g., absolute versus relative thresholds, geometric criteria). This reduces the amount of data that needs to be shown on a plot and allows one to see more clearly the similarities between methods based on similar criteria being met.

Ultimately, Tier 1 (and the overview paper) will aim to quantify the uncertainty in AR climatology (e.g., frequency, duration, and intensity), precipitation, and related impacts that arise because of different AR tracking methods. This will provide a baseline for comparison to results of Tier 2 analyses, which are based on climate model runs, and an assessment of how AR climatology and AR-related impacts may change in the future.

2.4 Tier 2 summary and status

Tier 2 will begin in the summer of 2018 and will focus on climate change as its first subtopic. The purpose of the second day of the workshop was to flesh out the details and create a timeline for Tier 2 projects. In order to reduce the burden of creating catalogues from a variety of data sets all at once, each subtopic will be initiated sequentially, starting with climate change and followed by reanalysis sensitivity. The goal of Tier 2 is to provide data and analysis delving into topical science questions important to research groups and stakeholders.

Each of the two Tier 2 subtopics (ARs in future climate and reanalysis sensitivity of ARs) was discussed at length both in plenary and in breakout groups. The high-resolution climate change project elicited the greatest interest. To be led by Ashley Payne, this study will analyze the LBNL high-resolution fvCAM5 simulations of a recent historical period and the end of century under RCP8.5 forcing conditions. Another highly motivated group discussed a similar analysis of the CMIP5 models. This study of lower-resolution models will have the advantage of using multi-model ensembles (at least 8 and perhaps 14 different models) and will be led by Travis O'Brien. A related multi-model analysis of simulations with prescribed SST (AMIP) and fully coupled (historical) CMIP5 models will explore the role of atmosphere-ocean coupling and will tentatively be led by Aneesh Subramanian. Another tentative modeling project would explore the role of horizontal resolution by an analysis of the 2º, 1º and 0.25º LBNL fvCAM5 simulations. Finally, an expansion of the Tier 1 project will apply the tracking algorithms to a variety of different reanalysis products to assess the source of structural uncertainty in describing real-world AR statistics. Progress on the modeling projects is ongoing. Michael Wehner has completed calculating high-frequency integrated water vapor fluxes and precipitable water products for both the high-resolution fvCAM5 and the CMIP5 models. These data are available via the LBNL data portal.

2.5 Future plans and timeline

ARTMIP participants will lead scientific papers focusing on their areas of interests. The ARTMIP committee will propose to organize a special collection of ARTMIP-related publications to be published in AGU journals. We will continue to solicit catalogues for the Tier 2 subtopics according to the timeline (Appendix A2), and will publish catalogues for the various tiers as they complete. Our hope is to provide the community with a data repository and easy access to all ARTMIP catalogues with the goal of encouraging AR science advancement.

2.6 Gaps and research priorities for ARTMIP

From the discussion of Tier 1 and Tier 2 activities, workshop participants identified gaps that need to be filled and research priorities that will enhance the impacts of ARTMIP for AR science and societal benefits. These are summarized briefly below.

Gap: As discussed in Section 2.4, different AR tracking methods can produce significant variations in AR statistics such as AR frequency, intensity, and duration. It will be necessary to understand the apparent spread in the statistics and characteristics of ARs among AR algorithms and allow recommendations to be made for what AR algorithms are more appropriate for addressing different scientific questions.

Research Priorities: It will be critically important to define a categorization of AR algorithms according to the regions within individual ARs that the algorithms tend to identify as part of the ARs. AR identification and tracking methods should also be grouped into "clusters" on the basis of key algorithm criteria such as geometric requirements. It must also be clearly communicated that these expert-developed methods were developed to answer different questions such as the impacts of ARs on flooding, so it makes sense that the AR statistics vary depending on the scientific goals. This information adds context to ARTMIP results that might otherwise be perceived as having too much spread, and hence large uncertainty.

Gap: There is still a major outstanding question of whether our quantitative—and possibly qualitative—understanding of how ARs may change in the future depends on the details of the algorithm used. Such potential sensitivities may influence both the detection and attribution of AR changes and understanding of the thermodynamical and dynamical mechanisms for the AR changes and the cascading AR impacts.

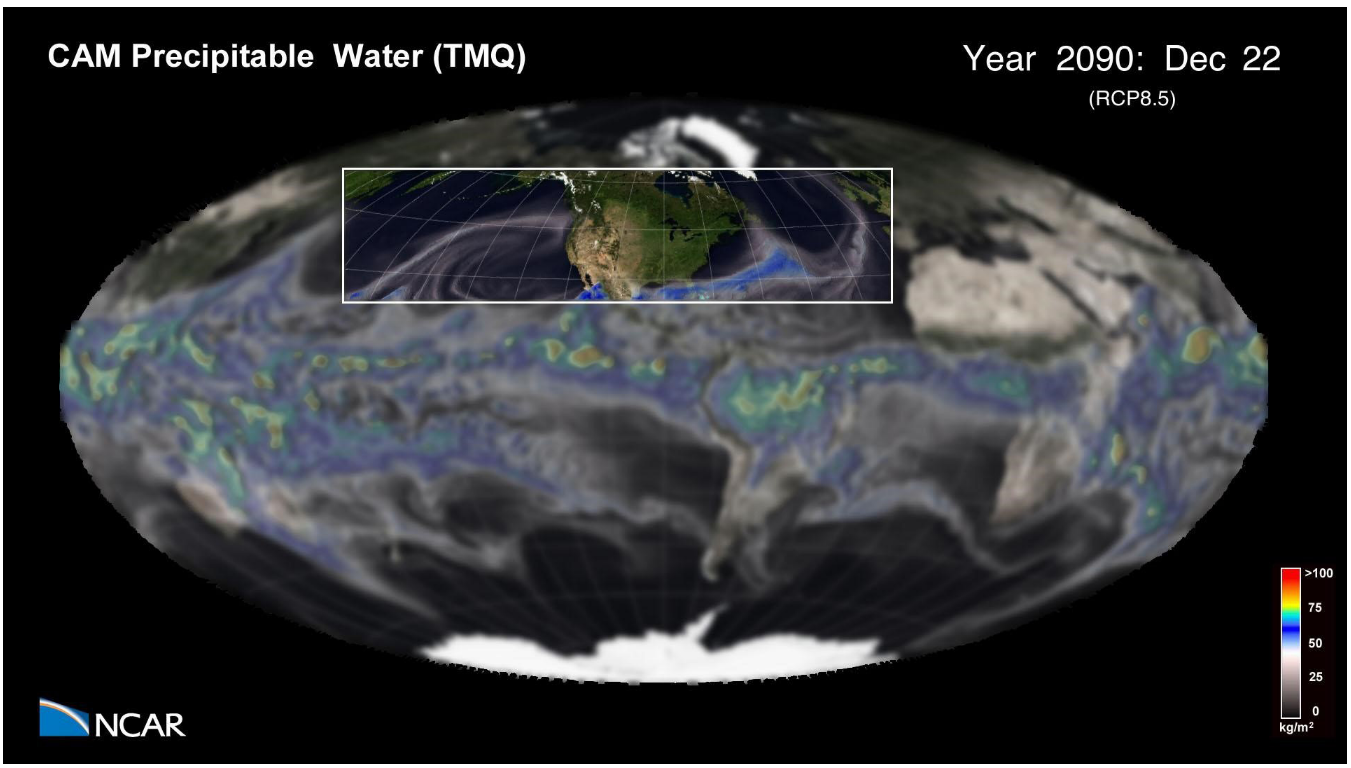
Research Priorities: Active multi-institutional participation and collaboration in the Tier 2 experiments analyzing ARs in high-resolution climate simulations and CMIP5/6 simulations is key to resolving whether and how our understanding of ARs and their changes in the future may depend on the tracking algorithm, model resolution, and other uncertainty factors. Such analyses can be placed in the context of existing literature on projected changes in ARs in different ocean basins to add insights on the mechanisms for the future changes.

Gap: AR identification and tracking methods choose different IVT thresholds (either as an absolute value or anomaly/percentile) to act as a lower bound for identifying ARs. Hence, AR intensity for each method invariably scales as a function of where this lower bound is set. A major challenge is communicating to stakeholders how the choice of this lower bound, and other criteria, affect the impacts associated with ARs identified by different methods.

Research Priorities: The definition of AR intensity can be informed and refined by ARTMIP, but different methods will likely persist. This may have important implications for scaling ARs so potential inconsistency or uncertainty should be resolved. Ultimately, ARTMIP should focus on developing novel ways of communicating uncertainty information that arises from differences in method criteria to stakeholders and the public.

Gap: There is a need for systematic analysis and quantification of uncertainties, and the implication of those uncertainties, in AR climatology and AR-related impacts that arise from the use of different AR tracking methods throughout the literature. These analyses, which will inform both the science and application communities, can only be facilitated by the maintenance of an organized data set based on these methods. Since the Tier 1 results were discussed in an overview paper (Shields et al. 2018), the ARTMIP data set has already grown with more participation using a broader set of AR detection algorithms.

Research Priorities: Catalogues should be made available to the scientific community to address science questions that will inevitably arise from ARTMIP analyses. Data management and storage is required for Tier 1 and 2 catalogues and will be housed on the Climate Data Gateway at NCAR as the project progresses.



3. BEYOND ARTMIP: EMERGING NEEDS AND OPPORTUNITIES FOR ATMOSPHERIC RIVERS RESEARCH

The plenary session on the second day of the workshop was dedicated to a discussion of emerging needs and opportunities for AR research. A number of topics were discussed including the needs of both weather and climate communities, observational research, and AR impacts.

3.1 AR forecasting and observations

Ways to improve AR forecasts were discussed, including topics such as identifying the metrics necessary for longer lead times, improved forecast systems, the need for a common vocabulary, and better ways to communicate AR impacts to the public. Improving forecast systems is tied to better use of data and tools. Important science questions to be addressed include: Can we understand AR metrics more fully with better use of satellite and in situ data? How can data from field campaigns such as Calwater (*Ralph et al.* 2016), Atmospheric Radiation Measurement (ARM) Cloud Aerosol Precipitation Experiment (ACAPEX) (https://www.arm.gov/publications/programdocs/doe-sc-arm-16-012.pdf), and Atmospheric River Reconnaisance be used in combination with modeling experiments to improve forecast and simulation of ARs and hydrologic extremes? With increasing model resolution supported by high-performance computing and numerical modeling advances such as regional refinement using unstructured grids, there is a need for high-resolution data sets and systematic evaluation of the impacts of model resolution on forecast and simulation of ARs and extreme precipitation. How can we better use observation data and metrics to help evaluate model biases and develop and apply process-level diagnostics?

Results from ARTMIP will provide enormous value to end users by providing understanding and estimates of uncertainty associated with AR climatology, AR projections, and AR impacts that arise from diverse algorithms used to identify and track ARs. Economic growth and vitality of the western United States depends on effective and efficient use of the existing water supplies. In the Pacific Northwest, the Columbia and Willamette Rivers have been controlled and diverted to support navigation interests in the region but are the main water supply for nearly 8 million people. The state of California, one of the largest in the west, developed its water resources with the California State Water Project, plus water from the Colorado River, to meet its water supply demands for more nearly 39 million people. Six other states and Mexico help California harness the Colorado River, which is the source of water for nearly 40 million people (*Water Education Foundation, Colorado River Basin,* 2017 http://www.watereducation.org/colorado-river-project). All of these rivers are driven by precipitation in the form of rain and snow that have a moisture source from the Pacific Ocean. This moisture is driven inland by storm systems frequently linked to atmospheric rivers that have a wide range of intensity and storm tracks from year to year.

The ability to make optimum use of the existing water supplies greatly depends on our ability to forecast the track of these storm systems and associated moisture throughout the year. An ability to forecast the intensity of the wet season, up to two years in advance, would greatly improve water managers' abilities to better operate control structures on the primary rivers and provide a more consistent water supply for multiple uses from recreation, navigation, agriculture, and human consumption, to name just a few. Studies conducted by the Bureau of Reclamation in 2007 recommended the Bureau to provide a "24-Month Study" of expected water availability in the Colorado River over a running 24 months in advance (*Bureau of Reclamation Report: Colorado River Interim Guidelines for Lower Basin Shortages and Coordinated Operations for Lake Powell and Lake Mead Final Environmental Impact Statement,* 2007).

The need to provide improved subseasonal to seasonal forecasting has been recognized by many and has been supported by Congress through the enactment of the Weather Research and Forecasting Innovation Act of 2017. The act directs the National Weather Service to improve forecasts, which implies doing research on many weather phenomena, including ARs, all with the primary goal of improving forecasting of precipitation for use by water supply users in the west. A significant effort was put forth by the Western States Water Council, a consortium of 18 governor-appointed member states with California as a major contributor, to support this bill as they developed position papers and documented the need and requirements for this information (*Western States Water Council: Position Number 399*, April 14, 2017).

3.2 AR climate research

Predictability on subseasonal, decadal, and longer timescales were highlighted as key emerging needs for AR climate research. What are the meteorological conditions needed for AR initiation and what factors influence AR evolution? What are the sources of predictability of ARs at weather, subseasonal-to-seasonal, and seasonal-to-decadal timescales? ARs in various ocean basins may be influenced differently by different modes of variability such as Madden-Julian Oscillation, North Atlantic Oscillation, and El Niño Southern Oscillation. How may the predictability of ARs vary in the North and South Pacific, North Atlantic, and Southern Ocean, and what factors are responsible for the potential differences in predictability? Numerical experiments using a hierarchy of models with different complexities and model resolutions may provide important insights to these questions. Given the large variability of ARs, ensemble modeling will be critical for extracting signals from noise to understand the sources of predictability of AR and extreme precipitation.

Delving deeper into the hydrological aspects associated with ARs was also noted. In addition to diagnosing and analyzing the relationship between AR and precipitation, quantities such as soil moisture, snowpack, and streamflow are all critical components of the local water cycle. Depending on the surface elevation and distance from the coast, different fractions of precipitation from ARs may be accumulated as snowpack during winter and spring or contribute to surface and subsurface runoff. Snowmelt water from snowpack that accumulates during ARs could constitute an important source of water for runoff in the summer and mitigate drought or water deficit in arid regions. How will surface and subsurface hydrology be affected by ARs on different climatic timescales? What characteristics of ARs are most predictive of extreme precipitation and floods? Extending the concept of AR scaling (*Ralph et al.* 2018a), can different categories of ARs be quantitatively related to water supply versus hazards such as flooding, landslide, and extreme surface winds (*Waliser and Guan* 2017)?

Polar ARs, in particular those connected to extreme events such as anomalous snowfall or rapid glacier mass loss, are relatively new areas for AR research. A few anomalous regional precipitation events can slow down Antarctic mass loss. In 2009, snowfall in Antarctica was unprecedented since 1979, leading to the largest surface mass balance anomaly during the last 60 years (*Lenaerts et al.* 2013). Isentropic analysis of a high-accumulation event in 2009 indicates the tropical origin of the precipitation event and shows the structure of an AR penetrating into Antarctica (*Gorodetskaya et al.* 2014). However, precipitation associated with ARs can cause snow accumulation or snow melt, depending on the AR temperature. Although ARs do reach far poleward, not all ARs make landfall in the Arctic and Antarctica, so there is a need to understand the mechanisms for polar ARs, such as the role of SST, evaporation, clouds, and atmospheric dynamics. There is also a need for more observations to understand the characteristics of enhanced moisture transport by ARs into the polar regions and to evaluate global reanalysis of AR integrated vapor transport.

Last but not least, ARs in the paleoclimate record were also discussed as an area to be more fully explored. The wetter-than-present conditions in southwestern North America during the Last Glacial Maximum was attributed to precipitation increase due largely to atmospheric rivers associated with decreases in sea-level pressure across the eastern North Pacific (Lora et al. 2017). Paleoclimate reconstructions and earth system models are useful tools for analysis of precipitation and AR changes in paleoclimate conditions. Understanding and identifying ARs in the historical record and their meteorological context could emerge as a key to understanding future ARs.

3.3 AR impacts

In both weather and climate science, the how, why, and where of AR landfall ultimately leads to impacts for local communities. More research is needed to quantify the impacts of AR and improving predictions of such impacts. While the impacts of ARs on hydrologic extremes such as flooding and mitigation of droughts have been explored, ARs are also important causes for landslide, strong winds, and coastal processes as extreme precipitation from ARs may generate extreme runoff that increases soil erosion and nutrient transport to the coast. The various direct impacts of ARs may lead to cascading effects on water resources and energy production.

Data analysis and modeling will be useful to quantify and predict AR impacts, but improving communication of AR impacts is also important to improve societal response. Communication geared towards, for example, the general public or water resource managers is critically important. Impacts due to extreme precipitation and flooding, wind damage, and/or prolonged exposure to AR conditions can have economic consequences. The group discussed ways to better communicate AR metrics, making them understandable outside of the AR research community. Taking a more interdisciplinary approach and pairing with different scientific communities, such as hydrologists, was also seen as a needed activity.

3.4 Gaps and research priorities beyond ARTMIP

In summary, workshop participants identified important gaps and research priorities beyond ARTMIP to advance understanding, forecasting, and modeling of ARs and their impacts, which are briefly discussed below.

Gap: Understanding and characterizing uncertainties is important to translate science to actions. Uncertainties in AR impacts, in part defined by ARTMIP analysis, can be communicated to stakeholders such as local governments and water managers.

Research Priorities: ARTMIP data and analysis can be applied to impacts such as extreme precipitation and runoff. The impacts metrics should be given a range of uncertainty given the diverse set of algorithms and metrics available in the literature. Stakeholders need to discern which algorithms metrics are most appropriate for them. ARTMIP can provide guidance.

Gap: Beyond ARTMIP, more research is needed to improve forecasting of ARs for a broad range of lead times, develop metrics and diagnostics to understand and quantify model forecast and simulation skill, improve understanding of the sources of predictability of ARs from subseasonal to multi-decadal timescales, and improve understanding and communication of ARs and their impacts geared towards the general public and stakeholders in need of AR forecast and projections.

Research Priorities: ARTMIP provides an important framework for systematic analysis of ARs and their regional and global impacts, focusing on uncertainty associated with AR tracking methods. This should open the door to further advance AR science to bridge the gap between research and societal needs. Improving forecasting and modeling of ARs and their impacts requires more observations because the life cycle of ARs occurs mainly over the oceans where satellite data and global reanalyses, each with their own uncertainties, are the primary source of information. Observations from field studies can play an important role in providing data sets for evaluating satellite data and global reanalyses and supporting studies of processes that influence AR development and evolution and their impacts on the water cycle. High-resolution data sets are also needed to support high-resolution modeling that is becoming more feasible with advances in high-performance computing and numerical methods such as regional refinement using unstructured grids. Numerical experiments using a hierarchy of models can be effective for studying AR predictability, AR impacts, and ARs in the paleoclimate records.

4. REFERENCES

American Meteorological Society: Atmospheric River, Glossary of Meteorology, available at: http://glossary.ametsoc.org/wiki/atmosphericriver (last access: 18 July 2018), 2017.

Bureau of Reclamation Report: Colorado River Interim Guidelines for Lower Basin Shortages and Coordinated Operations for Lake Powell and Lake Mead Final Environmental Impact Statement, 2007.

Dettinger, M. D., F. M. Ralph, T. Das, P. J. Neiman, and D. R. Cayan, 2011: Atmospheric Rivers, Floods and the Water Resources of California. *Water*, 3, 445–478, doi:10.3390/w3020445, http://dx.doi.org/10.3390/w3020445.

Dong, L., L.R. Leung, F. Song, and J. Lu, 2018: Roles of SST Forcing and Internal Atmospheric Variability in Winter Extreme Precipitation Variability Along the U.S. West Coast. *Journal of Climate*, 31, 8039-8058, doi:10.1175/jcli-d-18-0062.1.

Gao, Y., J. Lu, and L. R. Leung, 2016: Uncertainties in Projecting Future Changes in Atmospheric Rivers and their Impacts on Heavy Precipitation over Europe. *Journal of Climate*, 29, 6711–6726, doi:10.1175/jcli-d-16-0088.1, http://dx.doi.org/10.1175/JCLI-D-16-0088.1.

Gelaro, R., W. McCarty, M.J. Suárez, R. Todling, A. Molod, L. Takacs, C.A. Randles, A. Darmenov, M.G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A.M. da Silva, W. Gu, G. Kim, R. Koster, R. Lucchesi, D. Merkova, J.E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S.D. Schubert, M. Sienkiewicz, and B. Zhao, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

Gimeno, L., F. Dominguez, R. Nieto, R. Trigo, A. Drumond, C.J.C. Reason, A. Taschetto, A.M. Ramos, R. Kumar, J. Marengo, 2016: Major Mechanisms of Atmospheric Moisture Transport and Their Role in Extreme Precipitation Events. *Annual Review of Environment and Resources*, 41, 117–141, doi:10.1146/annurev-environ-110615-085558.

Gorodetskaya, I. V., M. Tsukernik, K. Claes, M. F. Ralph, W. D. Neff, and N. P. M. Van Lipzig, 2014: The role of atmospheric rivers in anomalous snow accumulation in East Antarctica. *Geophysical Research Letters*, 41, 6199–6206, doi:10.1002/2014gl060881, http://dx.doi.org/10.1002/2014gl060881.

Guan, B., and D. E. Waliser, 2015: Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies. *Journal of Geophysical Research–Atmospheres*, 120, 12514–12535, doi:10.1002/2015jd024257. http://dx.doi.org/10.1002/2015jd024257.

Hagos, S.M., L.R. Leung, J.-H. Yoon, J. Lu, and Y. Gao, 2016: A Projection of Changes in Landfalling Atmospheric River Frequency and Extreme Precipitation Over Western North America from the Large Ensemble CESM Simulations. *Geophysical Research Letters*, 43, 1357–1363, doi:10.1002/2015GL067392.

Lenaerts, J. T. M., E. van Meijgaard, M. R. van den Broeke, S. R. M. Ligtenberg, M. Horwath, and E. Isaksson: 2013: Recent snowfall anomalies in Dronning Maud Land, East Antarctica, in a historical and future climate perspective. *Geophysical Research Letters*, 40(11), 2684–2688, doi:10.1002/grl.50559.

Lora, J. M., J. L. Mitchell, C. Risi, and A. E. Tripati, 2017: North Pacific atmospheric rivers and their influence on western North America at the Last Glacial Maximum. *Geophysical Research Letters*, 44, 1051–1059, doi:10.1002/2016GL071541.

Lu, J., G. Chen, L.R. Leung, D.A. Burrows, Q. Yang, K. Sakaguchi, and S. Hagos, 2015: Toward the Dynamical Convergence on the Jet Stream in Aquaplanet AGCMs. *Journal of Climate*, 28, 6763-6782, doi:10.1175/jcli-d-14-00761.1.

Ralph, F.M., M.D. Dettinger, M.M. Cairns, T.J. Galarneau, and J. Eylander, 2018a: Defining "Atmospheric River": How the Glossary of Meteorology Helped Resolve a Debate. *Bulletin of the American Meteorological Society*, 99, 837–839, https://doi.org/10.1175/BAMS-D-17-0157.1.

Ralph, F.M., J.J. Rutz, J.M. Cordeira, M. Dettinger, M. Anderson, L. Schick, and C. Smallcomb, 2018b: A scale to characterize the strength and impacts of atmospheric rivers. *Bulletin of the American Meteorological Society*, in review.

Ralph, F.M., A.M. Wilson, T. Shulgina, B. Kawzenuk, S. Sellars, J.J. Rutz, M.A. Lamjiri, E.A. Barnes, A. Gershunov, B. Guan, K. Nardi, T. Osborne, and G.A. Wick, 2018c: Comparison of atmospheric river detection tools: How many atmospheric rivers hit northern California's Russian River Watershed? *Climate Dynamics*, in review.

Rutz, J. J., W. J. Steenburgh, and F. M. Ralph, 2014: Climatological Characteristics of Atmospheric Rivers and Their Inland Penetration over the Western United States. *Monthly Weather Review*, 142, 905–921, doi:10.1175/mwr-d-13-00168.1, http://dx.doi.org/10.1175/mwr-d-13-00168.1.

Shields, C. A., J.J. Rutz, L.-Y. Leung, F.M. Ralph, M. Wehner, B. Kawzenuk, J.M. Lora, E. McClenny, T. Osborne, A.E. Payne, P. Ullrich, A. Gershunov, N. Goldenson, B. Guan, Y. Qian, A.M. Ramos, C. Sarangi, S. Sellars, I. Gorodetskaya, K. Kashinath, V. Kurlin, K. Mahoney, G. Muszynski, R. Pierce, A.C. Subramanian, R. Tome, D. Waliser, D. Walton, G. Wick, A. Wilson, D. Lavers, Prabhat, A. Collow, H. Krishnan, G. Magnusdottir, and P. Nguyen, 2018: Atmospheric River Tracking Method Intercomparison Project (ARTMIP): project goals and experimental design, *Geoscientific Model Development,* 11, 2455-2474, doi:10.5194/gmd-11-2455-2018. http://dx.doi.org/10.5194/gmd-11-2455-2018.

U.S. DOE, 2018: Climate and Environmental Sciences Division Strategic Plan 2018–2023, DOE/SC–0192, U.S. Department of Energy Office of Science (science.energy.gov/~/media/ber/pdf/CESD-StratPlan-2018.pdf).

Waliser, D., and B. Guan, 2017: Extreme winds and precipitation during landfall of atmospheric rivers. *Nature Geoscience*, doi:10.1038/NGEO2894.

Western States Water Council, 2017: Position Number 399, April 14, 2017.

Wick, G.A., P.J. Neiman, F.M. Ralph, and T.M. Hamill, 2013: Evaluation of Forecasts of the Water Vapor Signature of Atmospheric Rivers in Operational Numerical Weather Prediction Models. *Weather Forecasting*, 28, 1337–1352, http://doi.org/10.1175/WAF-D-13-00025.1.

Yang, W., and G. Magnusdottir, 2017: Springtime extreme moisture transport into the Arctic and its impact on sea ice concentration. *Journal of Geophysical Research–Atmospheres*, 122, 5316–5329, doi:10.1002/2016jd026324. http://dx.doi.org/10.1002/2016JD026324.

Zhu, Y. and R.E. Newell, 1998: A Proposed Algorithm for Moisture Fluxes from Atmospheric Rivers. *Monthly Weather Review*, 126, 725–735, https://doi.org/10.1175/1520-0493(1998)126<0725:APAFMF>2.0.CO;2.

5. APPENDICES

Appendix A1

AGENDA

Atmospheric River Tracking Method Intercomparison Project (ARTMIP) Workshop Agenda
Gaithersburg, MD, April 23-24, 2018
Gaithersburg Marriott Washingtonian Center, Lakeside Ballroom

The goal of ARTMIP is to understand and quantify uncertainties in atmospheric river (AR) science based on choice of detection/tracking methodology.

http://www.cgd.ucar.edu/projects/artmip/

MONDAY, APRIL 23

7:45am – Breakfast

***Those with posters hang them during this time. Poster viewing during breaks. ***

8:30am – Welcome from DOE and remarks on programmatic needs (*Gary Geernaert, Renu Joseph*)

8:45am – Welcome and workshop introduction and goals (*Jon Rutz, Christine Shields*)

9:00am – Self introduction by ARTMIP participants

9:15am – Introduction and perspectives from other programs (*Christine Shields, moderator*)

[e.g., USGCRP/CEWEX: Jennifer Saleem-Arrigo; *NASA: Jared Entin*; NOAA: Daniel Barrie/Annarita Mariotti]

9:30am – Advances and gaps in AR science (*Jon Rutz, moderator*)

- 09:30 – 09:45: Marty Ralph: AR reconnaissance
- 09:45 – 10:00: Ruby Leung: Modeling and diagnostics of ARs
- 10:00 – 10:15: Alexandre Ramos: Atmospheric rivers research in Europe
- 10:15 – 10:30: Irina Gorodetskaya: ARs in polar regions

10:30 – 11:00 – Break

11:00am – ARTMIP experimental design, goals, and updates on the geoscientific model development overview paper (*Christine Shields*)

11:15am – Tier 1, MERRA-2 data comparison, overview paper (*Jon Rutz*)

11:30am – Current analysis by ARTMIP participants (*Ruby Leung, moderator*)

- 11:30 – 11:40 Ashley Payne (Duration)
- 11:40 – 11:50 Juan Lora (Intensity)
- 11:50 – 12:00 Bin Guan (Gleckler diagram)
- 12:00 – 12:10 Scott Sellar (Case study)
- 12:10 – 12:20 Marty Ralph (ARDT)
- 12:20 – 12:30 Swen Brands (European reanalysis) (*Remote*)

12:30pm – Lunch

1:30pm – Overview paper for Tier 1 discussion (*Jon Rutz, lead*)

A discussion of the preliminary results for the GRL paper, with a focus on:

- Understanding the AR climatologies emerging from analysis of the Tier 1 data
- Discussion of analysis presented before lunch
- Discussion on the identification and selection of uniform, global precipitation data sets for the GRL paper
- Assign needed analysis roles and discuss the timeline for the paper
- Do we need to adjust metrics?
- Can we state recommendations of algorithmic type yet?

2:30pm – Break/poster session

3:30pm – More on Tier 1 analyses and science questions (*Jon Rutz, lead*)

Discussion of other analyses the group and/or individuals should pursue based on the Tier 1 (i.e., 1980-2017) AR catalogs. (Volunteers to lead other papers, form collaborations.) Examples include:

- Regional analyses of precipitation,
- Climatology of ARs in polar regions
- Trends over that 37-year period.
- Other case studies

5:15pm – Adjourn

TUESDAY, APRIL 24

7:45am – Breakfast

8:30am – Tier 2 update (*Christine Shields, Michael Wehner lead*)

- Review Tier 2 topics and science questions: climate change, CMIP5, reanalysis
- Update on the status of Tier 2 data sets and availability
- Guidance on how the group should move forward

9:00am – Tier 2 discussion (*Christine Shields, Michael Wehner, Ruby Leung lead*)

The group discusses what is needed for Tier 2 analysis and identifies some topics that should be pursued as a part of ARTMIP

- Solicit reanalysis and CMIP5 interest and leads and participants
- Discussion on next steps: Products/models, input data, staging catalogues, etc.
- Develop timeline

10:00am – Break

11:00am – Emerging needs and opportunities (*Ruby Leung and Aneesh Subramanian, lead*)

12:00pm – Lunch

1:00pm – Breakout groups (*Determined by group, leads for Tier 2 topics*)

Unstructured time for groups to meet and discuss Tier 2 (or Tier 1) analyses related to precipitation, climate change, reanalyses, etc. This should be a free-flowing time where participants are encouraged to switch between groups and explore different ideas.

2:45pm – Break

3:15pm – Full group reconvenes and discusses key outcomes from breakout groups (*Christine Shields, lead*)

- Goals established and key analyses to pursue
- Data that is needed (and who will produce it)
- Roles within each group
- Next steps, timeline, future meetings, and workshops

5:00pm – Workshop ends

POSTER SESSION

"A Comparison of Modeled, Observed, Observation Corrected Precipitation Associated with Atmospheric River Enhanced Extreme Precipitation Events in the United States" (*Allison Collow)*

"Genesis, Pathways, and Terminations of Intense Global Water Vapor Transport in Association with Large-Scale Climate Patterns" (*Brian Kawzenuk*)

"Diversity in Detection Algorithms for Atmospheric Rivers: A Community Effort to Understand the Consequences (*Christine Shields*)

"Topological Data Analyses and Machine Learning for Detection, Classification, and Characterization of Atmospheric Rivers" (*Grzegorz Murzynski*)

"Assessing Uncertainty in Deep Learning Techniques that Identify Atmospheric Rivers in Climate Simulations" (*Travis O'Brien*)

"A Tracing Algorithm for Analyzing the Atmospheric River Life Cycles: Origins/Terminations, Lifetime, Intensity, and Propagation" (*Yang Zhou*)

"A Conditional Method to Detect Atmospheric Rivers on the Southwest Coast of South America" (*Maximiliano Viale*)

Appendix A2

MASTER ARTMIP TIMELINE

Date	Topic	Comment
May 2018	GMD final revisions	Finish
May 2018	Tier 1 Overview paper	Begin
May 2018 – October 2018	Tier 2 High-Res Climate Change catalogues	Complete and turn in catalogues
May 2018 – July 2018	Tier 2 CMIP5 acquire data	Begin, target July
May 2018 – December 2018	Tier 2 Reanalysis	Lead assign, start radiosonde comparison, reduce MERRA-2 data, establish timeline
July 2018	Tier 1 Overview paper	Draft
July 2018 – August 2018	Tier 2 CMIP5 process data	IVT, IWV, etc.
Sep 2018 – September 2019	Tier 2 CMIP6 data	Acquire/process, as available
Oct 2018 – December 2018	Tier 2 High-Res CC Overview paper	Analysis and draft
Jan 2019 – June 2019	Tier 2 CMIP5 catalogues	Complete and turn in catalogues
Jun 2019 – September 2019	Tier 2 CMIP5 Overview paper	Draft and submit
Jun 2019 – December 2019	Tier 2 CMIP6 catalogues	Complete and turn in catalogues
Sep 2019 – December 2019	Tier 2 CMIP5 versus CMIP6	Analysis
Dec 2019 – July 2020	Tier 2 Reanalysis catalogues	Choose products, compute group variables, begin catalogues
January 2020	Tier 2 CMIP5 versus 6 Paper	Draft

Appendix A3

Organizing Committee

Jonathan Rutz (NOAA)

Christine Shields (NCAR)

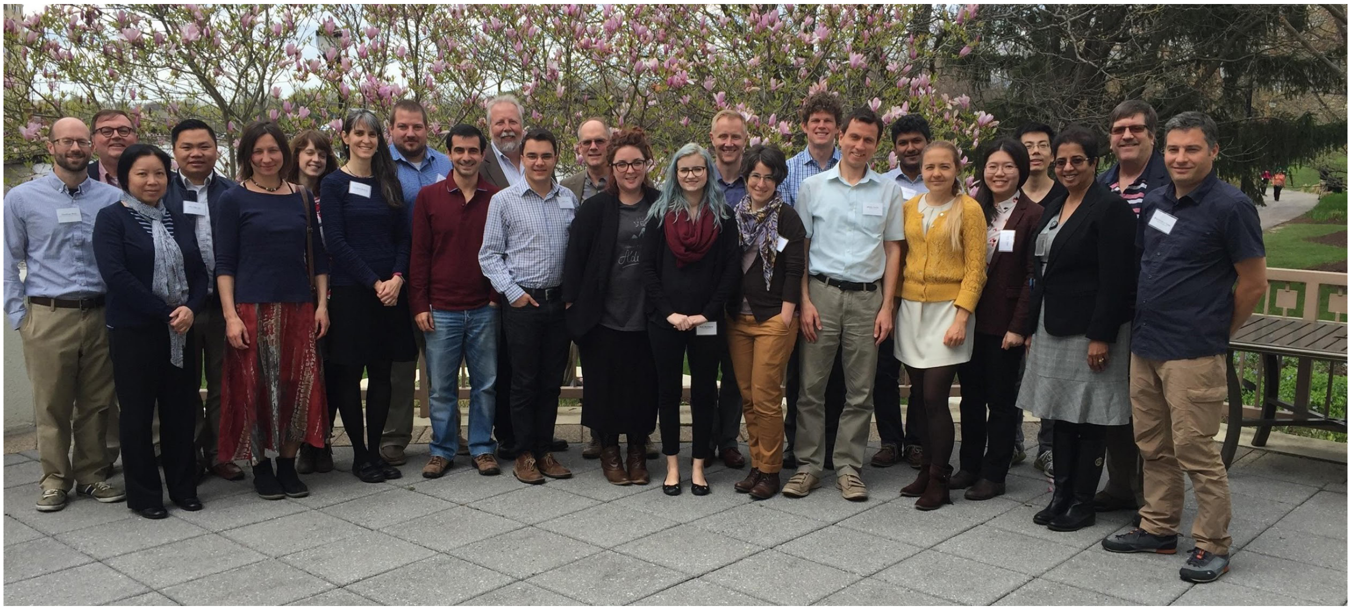
Ruby Leung (PNNL)

Michael Wehner (LBNL)

Marty Ralph (Scripps/CW3E)

Appendix A4

Workshop and Program Manager Participants



Left to Right: Jon Rutz, Roger Pierce, Ruby Leung, Phu Nguyen, Irina Gorodetskaya, Helen Griffith, Christine Shields, Brian Kawzenuk, Alexandre Ramos, Marty Ralph, Juan Lora, Gary Geernaert, Ashley Payne, Elizabeth Mclenny, Travis O'Brien, Naomi Goldenson, Daniel Walton, Vitaliy Kurlin, Aneesh Subramanian, Tamara Shulgina, Yang Zhou, Bin Guan, Renu Joesph, Michael Wehner, Maximilliano Viale. Not pictured (remote participation): Paul Ullrich, Swen Brands, Anna Wilson

Workshop Participants

Full Name	Company	Email
Renu Joseph	Department of Energy	renu.joseph@science.doe.gov
Swen Brands*	MeteoGalicia	swen.brands@gmail.com
Allison Collow	Universities Space Research Association	allison.collow@nasa.gov
Naomi Goldenson	University of California, Los Angeles	naomi.goldenson@gmail.com
Irina Gorodetskaya	University of Aveiro, Portugal	irina.gorodetskaya@ua.pt
Helen Griffith	University of Reading	h.v.griffith@pgr.reading.ac.uk
Bin Guan	Jet Propulsion Laboratory	bin.guan@jpl.nasa.gov
Brian Kawzenuk	University of California, San Diego	bkawzenuk@ucsd.edu
Vitaliy Kurlin	University of Liverpool	vitaliy.kurlin@gmail.com
Juan Lora	University of California, Los Angeles	jlora@ucla.edu
Ruby Leung	Pacific Northwest National Laboratory	Ruby.Leung@pnnl.gov
Kelly Mahoney	National Oceanic and Atmospheric Administration	kelly.mahoney@noaa.gov
Beth McClenny	University of California, Davis	eelliott@ucdavis.edu
Travis O'Brien	Lawrence Berkeley National Laboratory	taobrien@lbl.gov
Phu Nguyen	University of California, Irvine	ndphu@uci.edu
Ashley Payne	University of Michigan	aepayne@umich.edu

Roger Pierce	National Oceanic and Atmospheric Administration	roger.pierce@noaa.gov
Marty Ralph	University of California, San Diego	mralph@ucsd.edu
Alexandre Ramos	University of Lisbon	amramos@fc.ul.pt
Jonathan Rutz	National Oceanic and Atmospheric Administration	jonathan.rutz@noaa.gov
Christine Shields	National Center for Atmospheric Research	shields@ucar.edu
Tamara Shulgina	University of California, San Diego	tshulgina@ucsd.edu
Aneesh Subramanian	University of California, San Diego	acsubram@ucsd.edu
Paul Ullrich*	University of California, Davis	paullrich@ucdavis.edu
Maximiliano Viale	Instituto Argentino de Nivologia, Glaciologia Y Ciencias Ambientales	maxi@dgf.uchile.cl
Michael Wehner	Lawrence Berkeley National Laboratory	mfwehner@lbl.gov
Daniel Walton	University of California, Los Angeles	waltond@ucla.edu
Anna Wilson*	Scripps Institution of Oceanography	anna-m-wilson@ucsd.edu
Yang Zhou	Stony Brook University	yang.zhou.1@stonybrook.edu

*Remote participation

Appendix A5

ACRONYMS

ACAPEX . . . ARM Cloud Aerosol Precipitation Experiment

AGU . . . . . American Geophysical Union

AMIP. . . . . Atmospheric Model Intercomparison Project

AMS . . . . . American Meteorological Society

AR . . . . . . atmospheric river

ARM . . . . . Atmospheric Radiation Measurement

ARTMIP . . . Atmospheric River Tracking Method Intercomparison Project

CAM5 . . . . Community Atmosphere Model, version 5

CESD . . . . Climate and Environmental Sciences Division

CMIP. . . . . Coupled Model Intercomparison Project

DOD. . . . . U.S. Department of Defense

DOE . . . . . U.S. Department of Energy

DOI . . . . . U.S. Department of the Interior

EPA. . . . . . U.S. Environmental Protection Agency

IVT. . . . . . integrated vapor transport

IWCG . . . . Integrated Water Cycle Group

IWV . . . . . integrated water vapor

LBNL. . . . . Lawrence Berkeley National Laboratory

MERRA-2 . . Modern-Era Retrospective analysis for Research and Applications, version 2

NASA . . . . National Aeronautics and Space Administration

NCAR . . . . National Center for Atmospheric Research

NOAA . . . . National Oceanic and Atmospheric Administration

NSF . . . . . National Science Foundation

RCP . . . . . Representative Concentration Pathway

SST. . . . . . sea-surface temperature

USDA . . . . U.S. Department of Agriculture

USGCRP. . . U.S. Global Change Research Program

UTC . . . . . Coordinated Universal Time