In Situ Climate Modeling for Analyzing Extreme Weather Events

Soumya Dutta
Los Alamos National Laboratory
United States
sdutta@lanl.gov

Natalie Klein
Los Alamos National Laboratory
United States
neklein@lanl.gov

Li Tang
Los Alamos National Laboratory
United States
ltang@lanl.gov

Jonathan David Wolfe
Los Alamos National Laboratory
United States
jonbob@lanl.gov

Luke Van Roekel
Los Alamos National Laboratory
United States
lvanroekel@lanl.gov

James Joseph Benedict
Los Alamos National Laboratory
United States
benedict@lanl.gov

Ayan Biswas
Los Alamos National Laboratory
United States
ayan@lanl.gov

Earl Lawrence
Los Alamos National Laboratory
United States
earl@lanl.gov

Nathan Urban
Brookhaven National Laboratory
United States
nurban@bnl.gov

ABSTRACT

The study of many extreme weather events requires simulations with high spatiotemporal data that can grow in size quickly. Storing all the raw data from such a large-scale simulation for traditional post hoc analyses is soon going to be prohibitive as the data generation speed is outpacing the data storage capability in supercomputers. In situ analysis has emerged as a solution to this problem; data is analyzed when it is being produced, bypassing the slower disk input/output (I/O). In this work, we develop a new in situ analysis pathway for Energy Exascale Earth System Model (E3SM) and propose an algorithm for analyzing the impacts of sudden stratospheric warmings (SSWs), which can cause extreme cold temperature outbreaks at the surface, resulting in hazardous weather and disrupting many socioeconomic sectors. We detect SSWs and model the surface temperature data distributions in situ and show that post hoc analysis using the distribution models can predict the impact of SSWs in the continental United States.

CCS CONCEPTS

• Mathematics of computing → Probabilistic inference problems; **Distribution functions**; **• Human-centered computing → Visualization application domains**; **• Computing methodologies → Distributed algorithms**.

KEYWORDS

In situ analysis, generalized extreme value distribution modeling, climate simulation, visualization, high performance computing.

ISAV 2021, St. Louis, MO,

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/xxxx.xxxx

ACM Reference Format:
Soumya Dutta, Natalie Klein, Li Tang, Jonathan David Wolfe, Luke Van Roekel, James Joseph Benedict, Ayan Biswas, Earl Lawrence, and Nathan Urban. 2021. In Situ Climate Modeling for Analyzing Extreme Weather Events. In *Proceedings of ISAV 2021: In Situ Infrastructures for Enabling Extreme-scale Analysis and Visualization (ISAV 2021).* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/xxxx.xxxx

1 INTRODUCTION

With recent increases in computing capabilities, climate scientists can now study the dynamics of various physical phenomena using high-resolution computational models. But the data generated by such simulations is becoming prohibitively large, and due to slow disk I/O, it will not be viable to store all the data for post hoc analysis. To address this issue, climate scientists must move toward in situ analysis strategies in which data is analyzed in real-time while it is being produced in supercomputers, minimizing expensive disk I/O.

Many phenomena important to climate predictions occur at temporal frequencies that challenge our current data storage and access capabilities. Sudden Stratospheric Warmings (SSWs) [10, 21] are one example of such a process. SSWs are of critical interest for energy security as these events lead to extremely cold air outbreaks over the United States (referred to as “polar vortex events” in the media). Accurate diagnosis of SSWs requires access to the three-dimensional atmospheric variables at a high temporal frequency so that (1) SSWs can be detected precisely and (2) the impact of SSWs on surface temperature (TS) variations can be modeled robustly. Since it is unknown when an SSW will occur, we would need to run high-resolution climate models for a long duration, simulating hundreds of years and storing vast amounts of climate data for a post hoc analysis. As we move toward the era of exascale computing [15], this post hoc pipeline will not scale due to the bottleneck stemming from slow disk I/O and extreme data sizes. To enable the scientists to perform in situ climate analysis, we demonstrate a novel in situ analysis pathway for the atmosphere model (EAM) of the US DOE’s Energy Exascale Earth System Model (E3SM) [14, 18]. Our in situ infrastructure allows users to write their analysis routines in high-level, high-performance Julia language minimizing their programming effort. Using our in situ pathway, we study SSWs and analyze their impact on TS variations over the continental USA (CONUS). We propose an in situ analysis pipeline that detects SSWs using EAM data and produces generalized extreme value (GEV) distribution models of daily minimum TS values for both non-SSW and post-SSW cases. Our algorithm runs with E3SM and outputs the TS GEV model parameters, which use only a fraction of the raw data storage and can be analyzed post hoc for exploratory analysis. Therefore, our contributions are twofold:

(1) We develop a novel in situ analysis pathway for the atmosphere model (EAM) of E3SM which allows in situ execution of climate analysis scripts written in the high-level Julia language, reducing users' programming efforts significantly.
(2) We propose a new in situ algorithm to study the impact of SSW on surface temperature (TS) variations by modeling the TS values for non-SSW and post-SSW cases using probabilistic GEV models.

2 RELATED WORKS

Due to the ever-increasing gap between computing capabilities and I/O speeds of the supercomputers, in situ data analysis has gained significant attention over the past decade [5]. This has led to the emergence of several in situ frameworks, such as Ascent [22], ParaView Catalyst [17], SENSEI [28], ADIOS [27], and VisIt libSIM [31]. An image-based approach for in situ data reduction, called Cinema, was proposed by Ahrens et al. in [1]. Besides directly visualizing the data in situ, many in situ data analysis algorithms have also been proposed. One of the primary focuses of such algorithms is to achieve data reduction using methods such as compression [24–26], univariate sampling [7, 8, 30, 32], or information-theory based multivariate sampling [12]. To obtain a comprehensive overview of existing in situ infrastructures and formalized terminologies developed by the in situ research community, please refer to [5, 11].

Various statistical distribution-based data modeling schemes have also been used for performing in situ data modeling and analysis. In situ distribution-based data modeling has been explored recently as a means for data reduction that preserves the statistical features of the data [13, 33]. In situ copula-based distribution modeling and analysis of multivariate data has been proposed by Hazarika et al. [19]. In this work, we model the surface temperature values using GEV distribution models so that extreme behavior of surface temperatures can be captured accurately [20, 29]. We use SSW as an in situ trigger event to decide which GEV model (non-SSW or post-SSW) will be updated. This is similar in spirit to the in situ triggers proposed by Larsen et al. in their work [23].

3 IN SITU CLIMATE ANALYSIS

First, we briefly discuss SSW and GEV distributions before presenting our in situ algorithm for inferring the impact of SSWs on surface temperature variations. SSWs happen during the winter months (November through March) and are relatively rare; they occur, on average, about once every two winter seasons in the Arctic region [4]. Hence, the SSW detection algorithm will need to run for multiple years to achieve a sufficient sample size. In this work, the non-SSW and post-SSW GEV models are incrementally updated in situ over time, depending on whether an SSW was detected. Finally, the fitted GEV models for non-SSW and post-SSW cases are compared post hoc for understanding the influence of SSW on surface temperature.

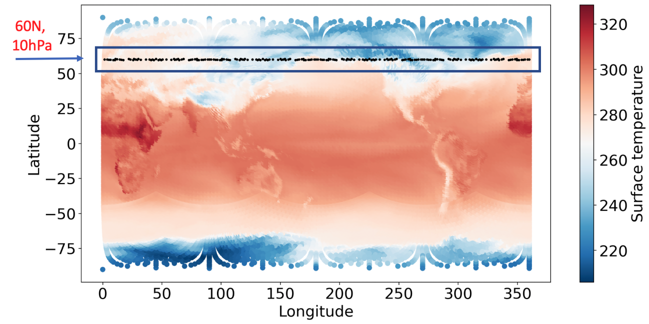

Figure 1: A scatter plot showing the data points required to estimate SSW. The black points (highlighted by dark blue box) located at 60°N and 10 hPa, satisfy the SSW definition and are used to compute SSW.

3.1 Sudden Stratospheric Warming (SSW)

As SSW is considered as one of the clear manifestations of the stratosphere-troposphere dynamic coupling, it is of interest to many climate scientists. It has been found that almost 46% of SSWs cause the splitting of the polar vortex [4, 10], and, in other cases, the polar vortex can shift, resulting in extreme cold temperatures at the surface. While several definitions of SSW exist in the literature, we have used the definition of SSW proposed by Andrews et al. [2] as a major midwinter warming that occurs when the daily *zonal mean zonal winds* at 60°N and 10 hPa (hectopascals) become easterly for at least 10 consecutive days between November and March [10]. In Figure 1, we show a scatter plot from EAM simulation data, colored by surface temperature, where the black points, highlighted by the dark blue box, depict the data points located at 60°N and 10 hPa that will be used to estimate the daily zonal mean zonal wind values.

3.2 Generalized Extreme Value (GEV) Model

Here, we want to study the distribution of extreme low surface temperatures and determine whether the distribution varies post-SSW (compared to non-SSW). Specifically, we seek to compare the probability of observing extreme low daily temperatures (below some threshold T_e) in general to the probability of such extremes following an SSW event. We use the Gumbel model, which is a member of the GEV family of probability distributions with probability density function

$$p(x;\mu,\beta) = \frac{1}{\beta}\exp\left\{-\left[\frac{x-\mu}{\beta} + \exp\left(-\frac{x-\mu}{\beta}\right)\right]\right\}. \quad (1)$$

The parameter μ is the mode of the distribution while the parameter β relates to the heavy-tailed behavior of the distribution. To estimate the posterior distribution of the model parameters, we use streaming variational inference to obtain a variational posterior approximation denoted $q(\mu,\beta)$ [9]. We chose this algorithm because it gives an estimate of uncertainty in the parameters and it

Algorithm 1: In situ algorithm for SSW-guided GEV modeling for analysis of extremes in surface temperature.

```
Input: N = No. of days from the current day that will be skipped
  before post-SSW GEV modeling starts when an SSW is detected.
Input: M = No. of consecutive days post-SSW GEV modeling is
  done.
Output: Fitted GEV models for both post-SSW and non-SSW cases.
for each time_step do
    Keep track of daily minimum TS values for GEV modeling.
    Counter C: Keeps track of consecutive negative daily zonal
      mean zonal wind values.
    if (current_time_step == end_of_day) then
        Compute the global zonal mean zonal wind with
          MPI:Reduction.
        if (SSW) then
            Update post-SSW GEV model using TS values with N
              days of forward time lag and continue GEV modeling
              for M consecutive days at each MPI process.
            Reset: C ← 0
        else
            Update non-SSW GEV model using TS values at each
              MPI process.
    else
        continue
```

applies to the in situ setting; it ingests one data point at a time to sequentially update q using variational inference (a computationally efficient approximation to full Bayesian inference). After estimating q using a stream of data, our post hoc analysis consists of predicting, as a summary of extreme cold temperatures, the probability of minimum daily temperatures below some extreme value threshold T_e (integrated over q via Monte Carlo integration). In addition, as an uncertainty metric, we compute the standard deviation of the estimated probabilities across samples from q.

3.3 SSW-guided In Situ Surface Temperature Modeling via GEV Distributions

The pseudo-code of our in situ analysis pipeline is shown in Algorithm 1. During the E3SM simulation, we access the EAM data at each time step and perform SSW-guided GEV modeling of surface temperature data. At each time step, we keep track of the daily minimum surface temperature (TS) values for each data point in CONUS and maintain a global variable that counts the number of consecutive days negative daily zonal mean zonal wind is detected between November and March. When a simulation time step marks the end of a day, we first compute the zonal mean zonal wind at each MPI process using the zonal velocity (U-velocity) variable of EAM. Since EAM mesh does not place data points exactly at 60°N and 10 hPa, we first filter out two layers of data points, which are above and below the 10 hPA level and fall within [59°N - 61°N] at each MPI rank. Then we linearly interpolate the zonal wind values to obtain values at 60°N and 10 hPa at each MPI rank. Finally, using an MPI reduction operation, we estimate the global zonal mean zonal wind value. If the value is negative, we increase the counter C by one. When the value of C becomes 10, i.e., the zonal mean zonal wind value is negative for 10 consecutive days, the parameters for the post-SSW GEV models for each spatial location in CONUS

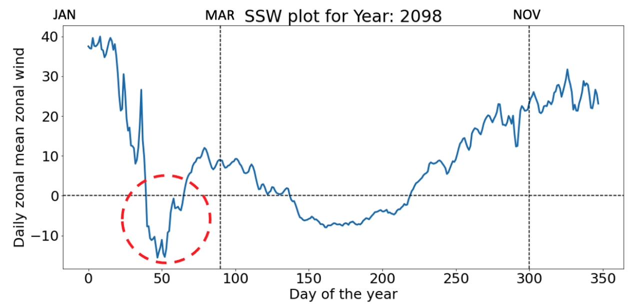

Figure 2: The Y-axis shows the daily zonal mean zonal wind values at 60°N, 10hPA and X-axis shows the day of the year. The plot indicates SSW detected for simulated year 2098 as highlighted by the red circle where the value of zonal mean zonal wind is negative for more than consecutive 10 days.

are updated. This update is not done immediately when SSW is detected since the impact of SSW on surface temperature is generally seen after a time lag. Therefore, we compute the day indices in which the post-SSW GEV models will be updated by adding N days forward time lag, and the post-SSW GEV model parameters are updated using data from M consecutive days. N and M in our pipeline are input parameters and scientists can use different time lags for testing their hypotheses. For all the other time steps, when SSW is not detected, we update a different set of GEV models per spatial location in CONUS using TS values which represent the non-SSW TS distributions. These GEV models, summarized by the approximate posterior distributions of their parameters, are stored on the disk for post hoc analysis and visualization.

4 ANALYSIS RESULTS

As a first step toward in situ analysis of various climate phenomena, we have conducted a long-term, climate-relevant simulation to generate sufficient data to develop and validate our analysis algorithms. To this end, we have run one realization of the Shared Socioeconomic Pathway (SSP) 585 scenario [16] with E3SM. This is an aggressive scenario that assumes the climate will experience an increase in radiative forcing of 8.5 W/m^2. This scenario is designed to provoke a strong model response and limit the influence of internal model variability. We use the standard E3SM V1 configuration with a 1° atmosphere and land (equivalent to 110 km at the equator), 0.5° river model (55 km), and an ocean and sea ice with mesh spacing varying between 60 km in the midlatitudes and 30 km at the equator and poles [18]. We started the simulation in year 2015 and simulated for 85 years in the future up to year 2100.

We ran the SSW detection algorithm for this offline data and found several simulated years when SSW was detected. We also found that the frequency of SSWs increased toward the later years, as was expected by the climate scientists. In Figure 2, we show a representative SSW plot for the year 2098 when an SSW event was detected. The Y-axis shows the zonal mean zonal wind at 60°N, 10 hPA. We observe a set of consecutive days (highlighted by the red circle) when the value of zonal mean zonal wind is negative, indicating the reversal of zonal mean winds and is identified as an

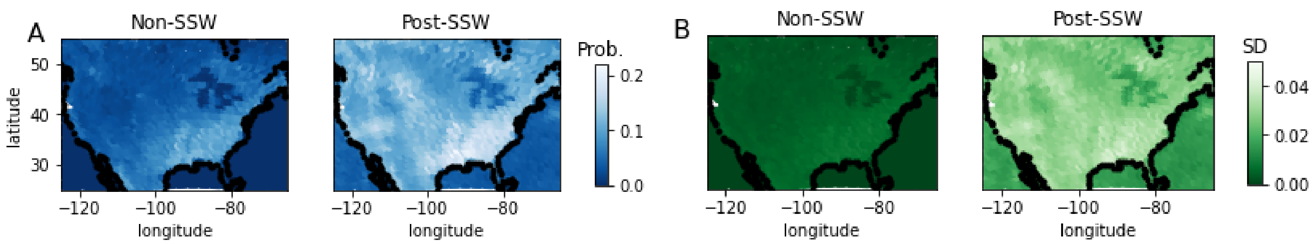

Figure 3: A) Predicted probabilities of observing extreme low temperatures in the CONUS (less than $10^{\circ}K$ below seasonal mean) in the non-SSW (left) and post-SSW (right) settings, indicating higher probabilities of extreme cold temperatures following an SSW event. B) Uncertainty in the predicted probabilities (represented by the standard deviation, or SD, across samples from the parameter posteriors). The uncertainty is higher for post-SSW events because SSW events are rarer.

SSW event. We can also see that there is another such time window when the zonal mean wind values are negative; however, since it does not happen between November and March, it does not qualify as a valid SSW event.

Next, we demonstrate the GEV modeling results using data from four simulated years in which SSW events were detected (specifically, 2083, 2085, 2095, and 2098). While the models were fit in a streaming fashion, we used an offline estimate of the mean temperature (computed over 30 simulated years of data) at each location and day of the year to first detrend the surface temperatures. Figure 3.A shows spatial maps of the estimated probability of observing extreme cold temperatures (defined as more than $10^{\circ}K$ below the average daily temperature) for the non-SSW (left) and post-SSW (right) regimes. It appears that the probability of extreme low temperatures is higher across most of the CONUS following SSW events. However, it is worth noting that because SSW events are fairly rare, the non-SSW models were fit using more observations than the post-SSW models. As shown in Figure 3.B, the uncertainty in the predicted probability of extreme events (represented using the standard deviation across Monte Carlo samples from the approximate posterior distributions of the parameters) is higher for the post-SSW model compared to the non-SSW model. While these results are preliminary, they demonstrate the capability for flexible in situ statistical modeling to answer scientific questions.

5 IN SITU STUDY

Our in situ integration enables the SSW-guided GEV modeling of temperature data generated by EAM following the streaming analysis shown in Algorithm 1. The E3SM code is developed in FORTRAN and EAM is the atmosphere module of E3SM that we have used. A schematic diagram of our in situ integration with EAM is shown in Figure 4. We aim to enable a Julia-based runtime environment for executing in situ analysis scripts using data generated from EAM. To the best of our knowledge, none of the existing in situ frameworks allow executing a Julia script in situ and so, in this work, we have developed a new in situ analysis pathway for EAM. Our in situ interface is lightweight and does not require recompilation of E3SM if the users want to change their analysis script. To access the EAM variables in situ, we have developed a FORTRAN-based in situ adapter that accesses the EAM data structures and reads the necessary variable arrays. Since Julia runtime currently only

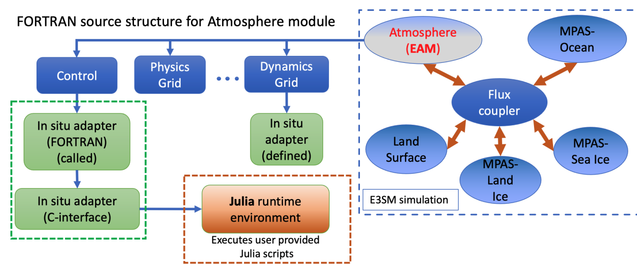

Figure 4: A schematic diagram of the in situ interface developed to integrate EAM with Julia run time environment.

supports C language embedding, we pass the data pointer first to an intermediate C-interface from FORTRAN, which then calls the Julia interface subroutine and passes the data pointer from C to the Julia runtime environment as shown in Figure 4. We invoke our FORTRAN in situ adapter from the Control module of EAM where the final version of the simulation data becomes available at each time step. Furthermore, since various climate analysis algorithms require accessing EAM data at different time frequencies, our in situ interface also provides access to E3SM's internal clock variables so that the users can call their in situ routines at the desired time-frequency. Note that the data from FORTRAN to C-interface undergo a one-time deep copy and then we pass the pointer from C-interface to the Julia environment.

5.1 In Situ Code Integration

Our in situ adapter also passes the E3SM MPI communicator to the Julia environment so that the algorithm developers can write their analysis routines using the same MPI communicator. The EAM data appears as Julia array objects in the Julia runtime environment and the users can write Julia routines using these objects directly. The motivation for using the Julia language comes from the fact that Julia offers the advantages of Python-like dynamic typed languages and has higher performance than other dynamic typed languages [6] and can be executed in GPU efficiently. Furthermore, the Julia modules can also be changed dynamically without recompilation of E3SM code due to Julia's JIT [3] feature. Even though, in this work, we only demonstrate the SSW-based analyses, our Julia-based in situ interface can be conveniently used for performing other in situ climate analyses simply by swapping the Julia script during runtime. Currently, our in situ adapter works only for the E3SM atmosphere model, but it can be easily extended for performing in situ analysis for other E3SM modules. In that case, we will have to modify our FORTRAN adapter to access data from other E3SM modules such as MPAS-O. Also, note that this in situ integration follows the direct integration strategy (i.e., the simulation and in situ analysis code use same computing resources), and once the in situ analysis is finished at each time step, the control goes back to E3SM and the simulation continues.

5.2 In Situ Results and Performance Evaluation

We ran our in situ algorithm with E3SM to measure the performance. The test case we used is the same configuration that was used to generate data discussed in Section 4. The in situ study was done on an HPC system, Grizzly, located at Los Alamos National Laboratory, consisting of 1490 computing nodes. Each node has 36 processor cores: $2 \times [E52695v4$ (i.e. Broadwell), 2.1GHz, 18 cores, 45MB cache], 128GB memory, and Intel OmniPath OP HFI, Single-port, PCIe-gen3x16 network interconnect.

As this E3SM case is optimized to run on 84 compute nodes (3024 cores), we also used 84 nodes, running E3SM for a duration of 3 simulated months. In Figure 5 (top) we show the SSW plot for the three-month in situ run. We observe that there is a small time window (marked by the red dotted circle) when the zonal mean zonal wind values are negative; however, since the number of such consecutive days is less than 10, it did not qualify for an SSW event. To validate the GEV modeling, we used the posterior mean in situ GEV parameters to analytically calculate the distributional mean, which represents the expected daily minimum temperature at each spatial point; the bottom image of Figure 5 shows this daily expected minimum temperature plot. The expected minimum temperatures exhibit spatial structure consistent with geography (including lower minimum temperatures in the North and the Rocky Mountain region, compared to other locations at similar latitudes).

We also found that the in situ processing time of our code at each time step is consistent and it takes on average 44.30 secs to simulate data for one day, whereas, if no in situ processing is done, then the atmosphere module (EAM) takes 40.046 seconds for simulating a day. Therefore our in situ processing is adding an overhead of 10.62%. Climate scientists in our team have acknowledged that this new Julia-based in situ analysis infrastructure for E3SM will be a valuable addition to their workflow and strongly encouraged us to continue the development and maintenance of this infrastructure for enabling in situ climate analysis using E3SM at scale. In the future, we plan to optimize our GEV estimation routine so that the in situ overhead comes down further to make our implementation more efficient.

6 CONCLUSION

We have presented a new in situ pathway for the E3SM atmosphere model which allows users to write their analysis scripts in Julia language. We believe that our in situ interface can reduce the programming efforts significantly for users without compromising the in situ performance. We also propose a statistical in situ analysis algorithm to study the impact of SSW and deploy it in situ to demonstrate the efficacy of our method. In the future, we plan on improving our in situ infrastructure and applying it to evaluate SSW-based analysis algorithms using various E3SM cases. We also intend to study other climate phenomena such as Madden-Julian Oscillation (MJO) in situ using high-resolution EAM data to derive enhanced understanding about such events, develop in situ inference capabilities, and study their impacts on human lives.

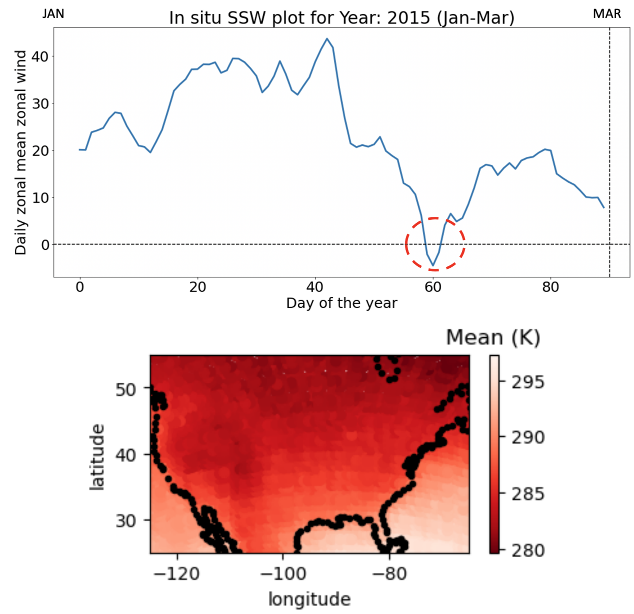

Figure 5: Top: SSW plot for the months of JAN-MAR generated from the in situ run and no SSW was detected. Bottom: Based on fitted GEV models from the in situ run, we show expected daily minimum temperatures in degrees Kelvin (K).

ACKNOWLEDGMENTS

Research presented in this paper was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project number $20200065DR$ and is released under LA-UR-21-28289. E3SM was obtained from the Energy Exascale Earth System Model project, sponsored by the U.S.Department of Energy, Office of Science, Office of Biological and Environmental Research. This research used resources provided by the Los Alamos National Laboratory Institutional Computing Program, which is supported by the U.S. Department of Energy National Nuclear Security Administration under Contract No. $89233218CNA000001$.

REFERENCES

[1] James Ahrens, Sebastien Jourdain, Patrick O'Leary, John Patchett, David H. Rogers, and Mark Petersen. 2014. An Image-Based Approach to Extreme Scale in Situ Visualization and Analysis. In *SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. 424–434. https://doi.org/10.1109/SC.2014.40

[2] David Andrews, Conway Leovy, and James Holton. 1987. *Middle atmosphere dynamics*. https://www.osti.gov/biblio/5936274

[3] John Aycock. 2003. A brief history of just-in-time. *ACM Computing Surveys (CSUR)* 35, 2 (2003), 97–113.
[4] Mark P. Baldwin, Blanca Ayarzagüena, Thomas Birner, Neal Butchart, Amy H. Butler, Andrew J. Charlton-Perez, Daniela I. V. Domeisen, Chaim I. Garfinkel, Hella Garny, Edwin P. Gerber, Michaela I. Hegglin, Ulrike Langematz, and Nicholas M. Pedatella. 2021. Sudden Stratospheric Warmings. *Reviews of Geophysics* 59, 1 (2021), e2020RG000708. https://doi.org/10.1029/2020RG000708
[5] Andrew C. Bauer, Hasan Abbasi, James Ahrens, Hank Childs, Berk Geveci, Scott Klasky, Kenneth Moreland, Patrick O'Leary, Venkatram Vishwanath, Brad Whitlock, and E. W. Bethel. 2016. In Situ Methods, Infrastructures, and Applications on High Performance Computing Platforms. *Computer Graphics Forum* (2016).
[6] Jeff Bezanson, Stefan Karpinski, Viral B Shah, and Alan Edelman. 2012. Julia: A fast dynamic language for technical computing. *arXiv preprint arXiv:1209.5145* (2012).
[7] Ayan Biswas, Soumya Dutta, Earl Lawrence, John Patchett, Jon C. Calhoun, and James Ahrens. 2020. Probabilistic Data-Driven Sampling via Multi-Criteria Importance Analysis. *IEEE Transactions on Visualization and Computer Graphics* (2020), 1–1. https://doi.org/10.1109/TVCG.2020.3006426
[8] Ayan Biswas, Soumya Dutta, Jesus Pulido, and James Ahrens. 2018. In Situ Data-Driven Adaptive Sampling for Large-Scale Simulation Data Summarization. In *Proceedings of the Workshop on In Situ Infrastructures for Enabling Extreme-Scale Analysis and Visualization (ISAV '18)*. Association for Computing Machinery, New York, NY, USA, 13–18. https://doi.org/10.1145/3281464.3281467
[9] Tamara Broderick, Nicholas Boyd, Andre Wibisono, Ashia C Wilson, and Michael I Jordan. 2013. Streaming variational bayes. *arXiv preprint arXiv:1307.6769* (2013).
[10] Andrew J. Charlton and Lorenzo M. Polvani. 2007. A New Look at Stratospheric Sudden Warmings. Part I: Climatology and Modeling Benchmarks. *Journal of Climate* 20, 3 (2007), 449 – 469. https://doi.org/10.1175/JCLI3996.1
[11] Hank Childs et al. 2020. A terminology for in situ visualization and analysis systems. *The International Journal of High Performance Computing Applications* 34, 6 (2020), 676–691. https://doi.org/10.1177/1094342020935991
[12] Soumya Dutta, Ayan Biswas, and James Ahrens. 2019. Multivariate Pointwise Information-Driven Data Sampling and Visualization. *Entropy* 21, 7 (2019). https://doi.org/10.3390/e21070699
[13] Soumya Dutta, Chun-Ming Chen, Gregory Heinlein, Han-Wei Shen, and Jen-Ping Chen. 2017. In Situ Distribution Guided Analysis and Visualization of Transonic Jet Engine Simulations. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 811–820. https://doi.org/10.1109/TVCG.2016.2598604
[14] E3SM Project. 2018. Energy Exascale Earth System Model (E3SM). [Computer Software] https://dx.doi.org/10.11578/E3SM/dc.20180418.36. https://doi.org/10.11578/E3SM/dc.20180418.36
[15] ECP (accessed August 16, 2021). ECP: Exascale Computing Project. https://www.exascaleproject.org/.
[16] Veronika Eyring, Sandrine Bony, Gerald A. Meehl, Catherine A. Senior, Bjorn Stevens, Ronald J. Stouffer, and Karl E. Taylor. 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development* 9, 5 (2016), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016
[17] Nathan Fabian, Kenneth Moreland, David Thompson, Andrew C. Bauer, Pat Marion, Berk Gevecik, Michel Rasquin, and Kenneth E. Jansen. 2011. The ParaView Coprocessing Library: A scalable, general purpose in situ visualization library. In *2011 IEEE Symposium on Large Data Analysis and Visualization*. 89–96. https://doi.org/10.1109/LDAV.2011.6092322
[18] Jean-Christophe Golaz et al. 2019. The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution. *Journal of Advances in Modeling Earth Systems* 11, 7 (2019), 2089–2129. https://doi.org/10.1029/2018MS001603 arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018MS001603
[19] Subhashis Hazarika, Soumya Dutta, Han-Wei Shen, and Jen-Ping Chen. 2019. CoDDA: A Flexible Copula-based Distribution Driven Analysis Framework for Large-Scale Multivariate Data. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (2019), 1214–1224. https://doi.org/10.1109/TVCG.2018.2864801
[20] Whitney K Huang, Michael L Stein, David J McInerney, Shanshan Sun, and Elisabeth J Moyer. 2016. Estimating changes in temperature extremes from millennial-scale climate simulations using generalized extreme value (GEV) distributions. *Advances in Statistical Climatology, Meteorology and Oceanography* 2, 1 (2016), 79–103.
[21] Junsu Kim, Seok-Woo Son, Edwin P. Gerber, and Hyo-Seok Park. 2017. Defining Sudden Stratospheric Warming in Climate Models: Accounting for Biases in Model Climatologies. *Journal of Climate* 30, 14 (2017), 5529 – 5546. https://doi.org/10.1175/JCLI-D-16-0465.1
[22] Matthew Larsen, James Ahrens, Utkarsh Ayachit, Eric Brugger, Hank Childs, Berk Geveci, and Cyrus Harrison. 2017. The ALPINE In Situ Infrastructure: Ascending from the Ashes of Strawman. In *Proceedings of the In Situ Infrastructures on Enabling Extreme-Scale Analysis and Visualization (ISAV'17)*. Association for Computing Machinery, New York, NY, USA, 42–46. https://doi.org/10.1145/3144769.3144778
[23] Matthew Larsen, Amy Woods, Nicole Marsaglia, Ayan Biswas, Soumya Dutta, Cyrus Harrison, and Hank Childs. 2018. A Flexible System for in Situ Triggers. In *Proceedings of the Workshop on In Situ Infrastructures for Enabling Extreme-Scale Analysis and Visualization (ISAV '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3281464.3281468
[24] Henry Lehmann and Bernhard Jung. 2014. In-situ multi-resolution and temporal data compression for visual exploration of large-scale scientific simulations. In *2014 IEEE 4th Symposium on Large Data Analysis and Visualization (LDAV)*. 51–58. https://doi.org/10.1109/LDAV.2014.7013204
[25] Xin Liang, Sheng Di, Dingwen Tao, Zizhong Chen, and Franck Cappello. 2018. An Efficient Transformation Scheme for Lossy Data Compression with Point-Wise Relative Error Bound. In *2018 IEEE International Conference on Cluster Computing (CLUSTER)*. 179–189. https://doi.org/10.1109/CLUSTER.2018.00036
[26] Peter Lindstrom. 2014. Fixed-Rate Compressed Floating-Point Arrays. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 2674–2683. https://doi.org/10.1109/TVCG.2014.2346458
[27] Jay F. Lofstead, Scott Klasky, Karsten Schwan, Norbert Podhorszki, and Chen Jin. 2008. Flexible IO and Integration for Scientific Codes through the Adaptable IO System (ADIOS). In *Proceedings of the 6th International Workshop on Challenges of Large Applications in Distributed Environments (CLADE '08)*. Association for Computing Machinery, New York, NY, USA, 15–24. https://doi.org/10.1145/1383529.1383533
[28] SENSEI 2021 (accessed August 16, 2021). SENSEI:Scalable in situ analysis and visualization. https://sensei-insitu.org/.
[29] Zhuo Wang, Yujing Jiang, Hui Wan, Jun Yan, and Xuebin Zhang. 2017. Detection and attribution of changes in extreme temperatures at regional scale. *Journal of Climate* 30, 17 (2017), 7035–7047.
[30] Tzu-Hsuan Wei, Soumya Dutta, and Han-Wei Shen. 2018. Information Guided Data Sampling and Recovery Using Bitmap Indexing. In *2018 IEEE Pacific Visualization Symposium (PacificVis)*. 56–65. https://doi.org/10.1109/PacificVis.2018.00016
[31] Brad Whitlock, Jean M. Favre, and Jeremy S. Meredith. 2011. Parallel in Situ Coupling of Simulation with a Fully Featured Visualization System. In *Proceedings of the 11th Eurographics Conference on Parallel Graphics and Visualization (EGPGV '11)*. Eurographics Association, 101–109. https://doi.org/10.2312/EGPGV/EGPGV11/101-109
[32] Jonathan Woodring, James Ahrens, J. Figg, Joanne Wendelberger, Salman Habib, and Katrin Heitmann. 2011. In-situ Sampling of a Large-scale Particle Simulation for Interactive Visualization and Analysis. In *Proceedings of the 13th Eurographics / IEEE - VGTC Conference on Visualization*. Eurographics Association, 1151–1160. https://doi.org/10.1111/j.1467-8659.2011.01964.x
[33] Yucong Chris Ye, Tyson Neuroth, Franz Sauer, Kwan-Liu Ma, Giulio Borghesi, Aditya Konduri, Hemanth Kolla, and Jacqueline Chen. 2016. In situ generated probability distribution functions for interactive post hoc visualization and analysis. In *2016 IEEE 6th Symposium on Large Data Analysis and Visualization (LDAV)*. 65–74. https://doi.org/10.1109/LDAV.2016.7874311