

# ROAMM: A Benchmark Dataset for Multimodal Human Attention Decoding and EEG-to-Text Modeling During Naturalistic Reading

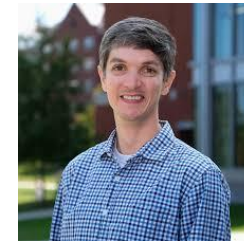
---

David Jangraw  
djangraw@uvm.edu

Haorui Sun, Ardyn Olszko, Niharika Singh, and David Jangraw



ICML 2026



# Motivation



## 1. Attention is not stable:

Mind-wandering (MW) may occupy 30–60% of daily life (Killingsworth & Gilbert, 2010)



## 2. MW is associated with degraded task performance and comprehension

(van Vugt & Broers, 2016; Smallwood et al., 2008; Unsworth & McMillan, 2013)



## 3. During reading, MW can disrupt linguistic processing even when outward reading behavior appears normal

(Reichle et al., 2010)

→ This confounds behavioral analyses and neural decoding



## 4. There is a need for datasets that capture reading under realistic conditions while providing attention-state annotations to support attention-aware modeling and evaluation

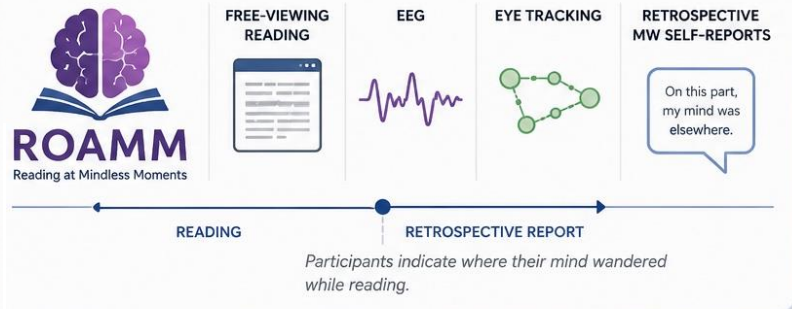
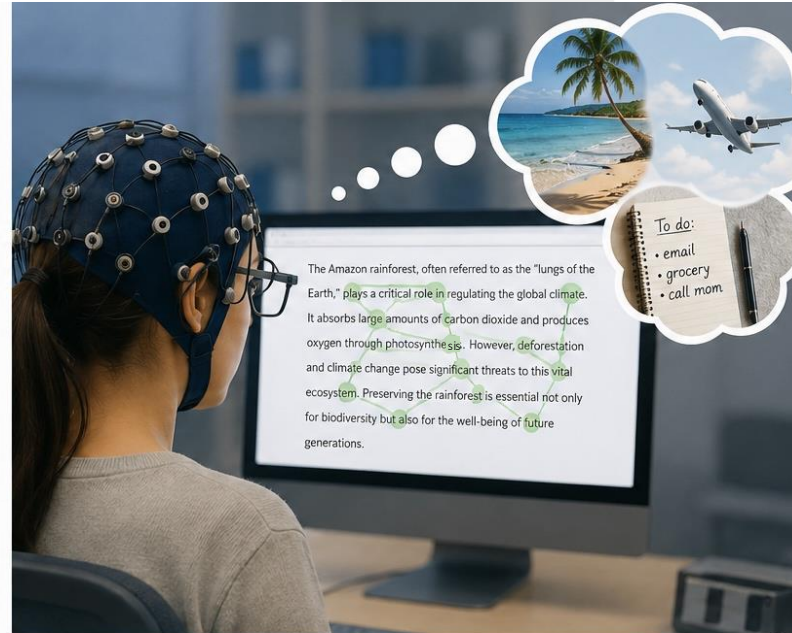
(Ho & Griffiths, 2022)



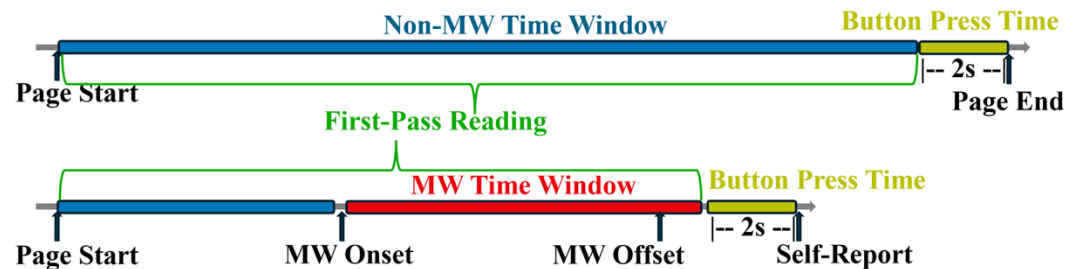
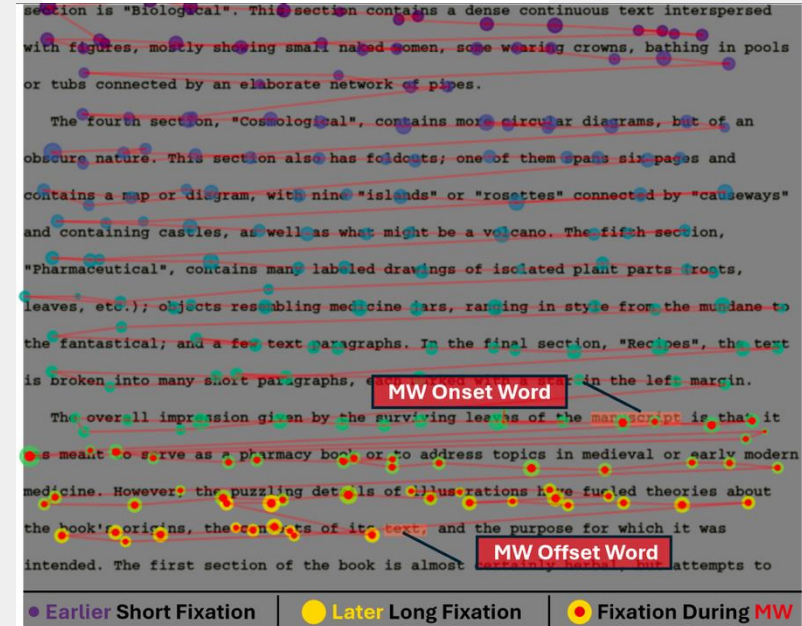
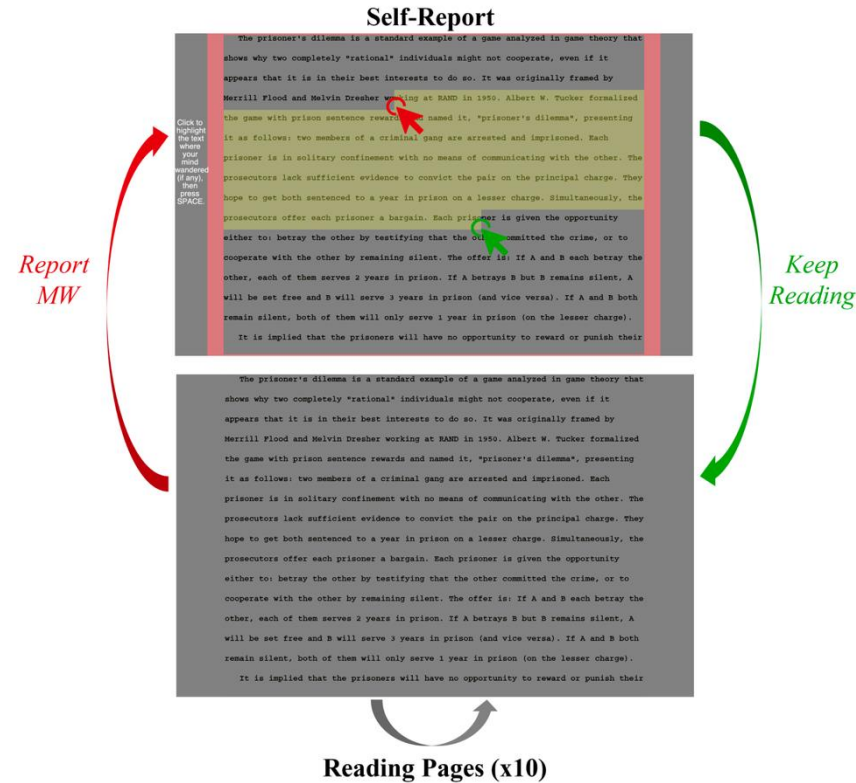
## 5. This is a longstanding challenge because self-reports or thought probes lack MW onset annotations.



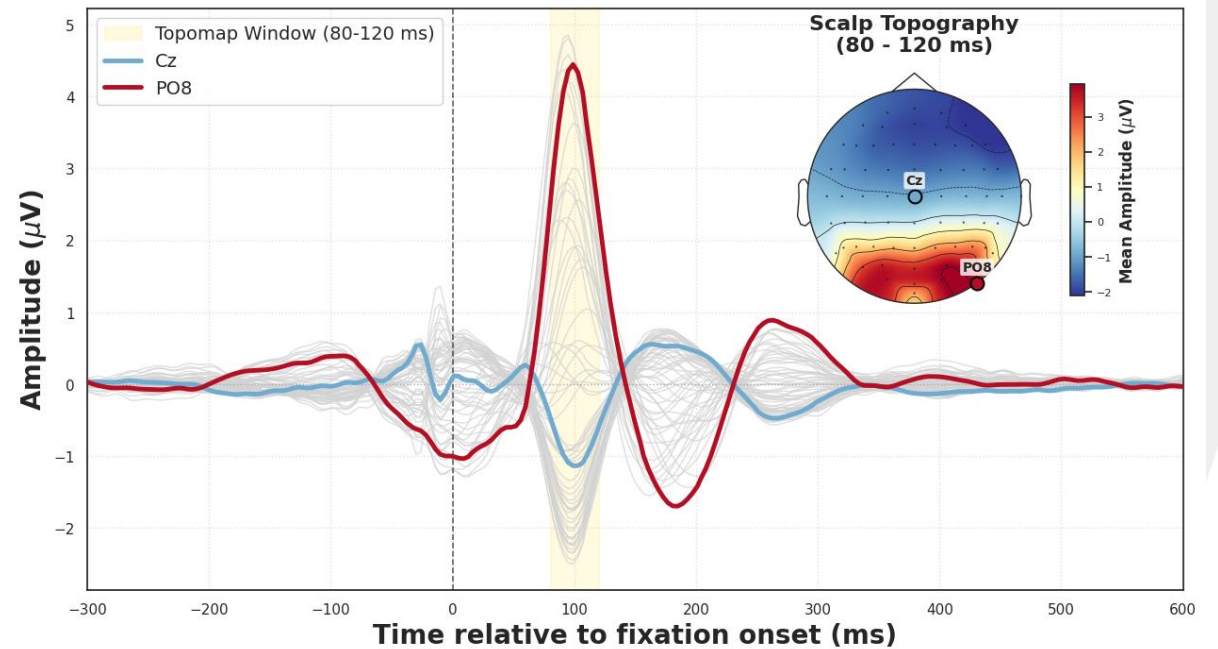
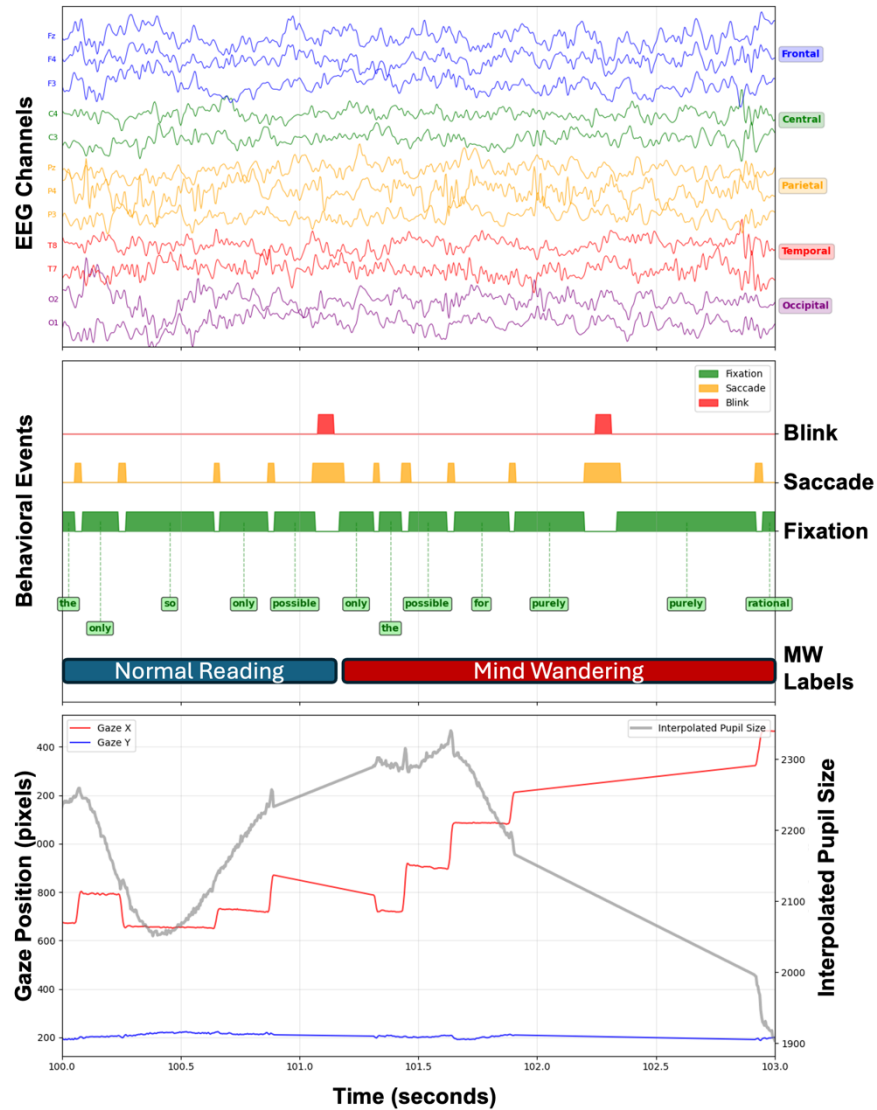
## 6. We present Reading at Mindless Moments (ROAMM), a dataset of free-viewing reading with EEG and eye tracking data that includes retrospective MW self-reports.



# ReMind Task: Free-Viewing Reading with Retrospective MW Reports



# ROAMM Dataset



# ROAMM Scope

## Large EEG + Eye Reading Dataset

Dataset	# Participants	Task	Length
EEGeyeNet	356	Symbol viewing	47 h
CG-EEGET	113	Gaze movement	12 h
EEGET-RSOD	38	Target search	–
ZuCo 1.0	12	Sentence reading	–
ZuCo 2.0	18	Sentence reading	–
<b>ROAMM</b>	44	<b>Page reading</b>	<b>50 h</b>

## Includes MW Annotations

DATA TYPE	#SAMPLES	TIME
TOTAL RECORDING	46,371,584	50H 18M
FIRST-PASS READING	26,691,014	28H 57M
MIND-WANDERING	2,042,601	2H 12M
FIXATIONS	38,316,711	41H 34M
SACCADES	9,323,085	10H 6M
BLINKS	2,287,380	2H 28M

## Full Column Annotations in Supplement

Column(s)	# Cols	Unit	Description
Fp1–O2	64	V	64 EEG channel voltage values using BioSemi 10–20 electrode labels.
time	1	s	Timestamp aligned to the synchronized EEG time base.
sfreq	1	Hz	Sampling frequency after synchronization/downsampling.
first_pass_reading	1	binary	Whether the sample belongs to first-pass reading.
page_num	1	index	Page number within the current article.
page_start, page_end	2	s	Start and end timestamps of the current page.
page_dur	1	s	Duration of the current page.
is_mw	1	binary	Whether the sample is labeled as mind-wandering.
mw_onset, mw_offset	2	s	Estimated onset and offset timestamps of the MW episode.
mw_dur	1	s	Duration of the MW episode.
run_num	1	index	Article/run number within the session.
story_name	1	text	Name of the Wikipedia article being read.
is_fix	1	binary	Whether the sample occurs during a fixation.
fix_[L/R]_eye	2	text	Eye associated with the fixation event.
fix_[L/R]_tStart, fix_[L/R]_tEnd	4	s	Start and end timestamps of the fixation.
fix_[L/R]_duration	2	ms	Duration of the fixation.
fix_[L/R]_xAvg, fix_[L/R]_yAvg	4	px	Average fixation position on the screen.
fix_[L/R]_pupilAvg	2	arbitrary unit	Average pupil area during the fixation.
fix_[L/R]_fixed_word	2	text	Word fixated during the event.
fix_[L/R]_fixed_word_key	2	UUID	Unique key linking the fixation to a word in the story file.
is_blink	1	binary	Whether the sample occurs during a blink.
blink_[L/R]_eye	2	text	Eye associated with the blink event.
blink_[L/R]_tStart, blink_[L/R]_tEnd	4	s	Start and end timestamps of the blink.
blink_[L/R]_duration	2	ms	Duration of the blink.
is_sacc	1	binary	Whether the sample occurs during a saccade.
sacc_[L/R]_eye	2	text	Eye associated with the saccade event.
sacc_[L/R]_tStart, sacc_[L/R]_tEnd	4	s	Start and end timestamps of the saccade.
sacc_[L/R]_duration	2	ms	Duration of the saccade.
sacc_[L/R]_xStart, sacc_[L/R]_yStart	4	px	Starting screen position of the saccade.
sacc_[L/R]_xEnd, sacc_[L/R]_yEnd	4	px	Ending screen position of the saccade.
sacc_[L/R]_ampDeg	2	degrees	Saccade amplitude in visual angle.
sacc_[L/R]_vPeak	2	deg/s	Peak saccade velocity.
tSample	1	s	Original eye-tracking sample timestamp.
LX, LY, RX, RY	4	px	Raw left/right gaze coordinates.
LPupil, RPupil	2	arbitrary unit	Raw left/right pupil measurements.
blink_interp_[LX/LY/RX/RY]	4	px	Blink-interpolated gaze coordinates.
blink_interp_[LPupil/RPupil]	2	arbitrary unit	Blink-interpolated pupil measurements.

# Mind-Wandering Detection

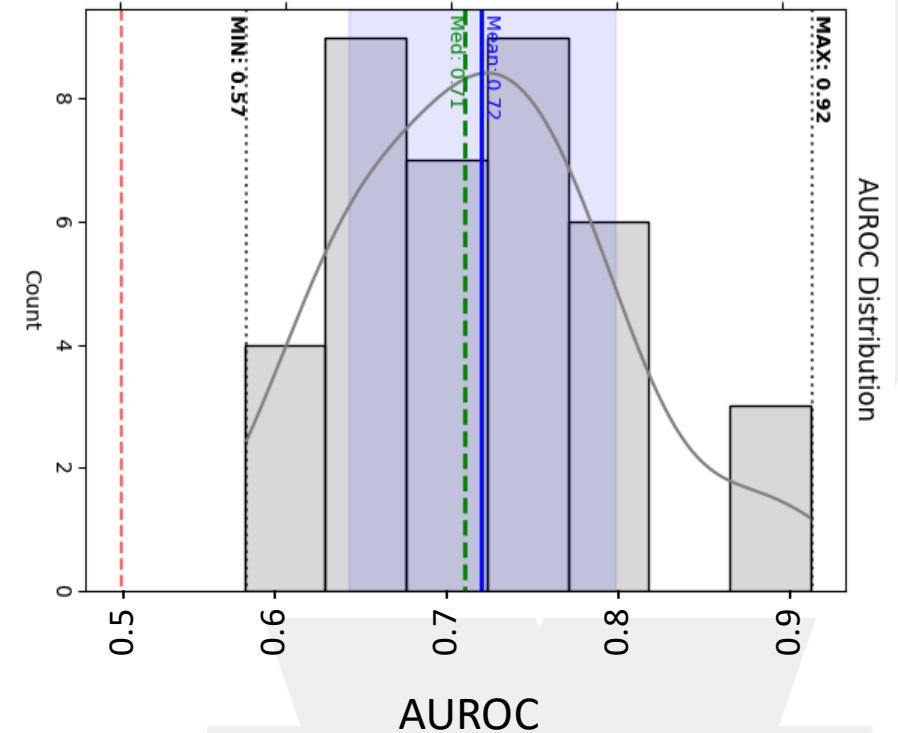
## Cross-Subject Models

Model	AUROC		
	ET	EEG	ET+EEG
LogReg	<b>0.592</b>	<b>0.569</b>	<b>0.599</b>
Linear SVC	<b>0.592</b>	<b>0.566</b>	<b>0.599</b>
Gradient Boosting	<b>0.588</b>	<b>0.572</b>	<b>0.587</b>
RBF SVC	0.586	0.569	0.600
kNN	<b>0.549</b>	0.544	<b>0.557</b>
MLP	<b>0.582</b>	<b>0.558</b>	<b>0.560</b>
Random Forest	<b>0.565</b>	<b>0.590</b>	<b>0.609</b>

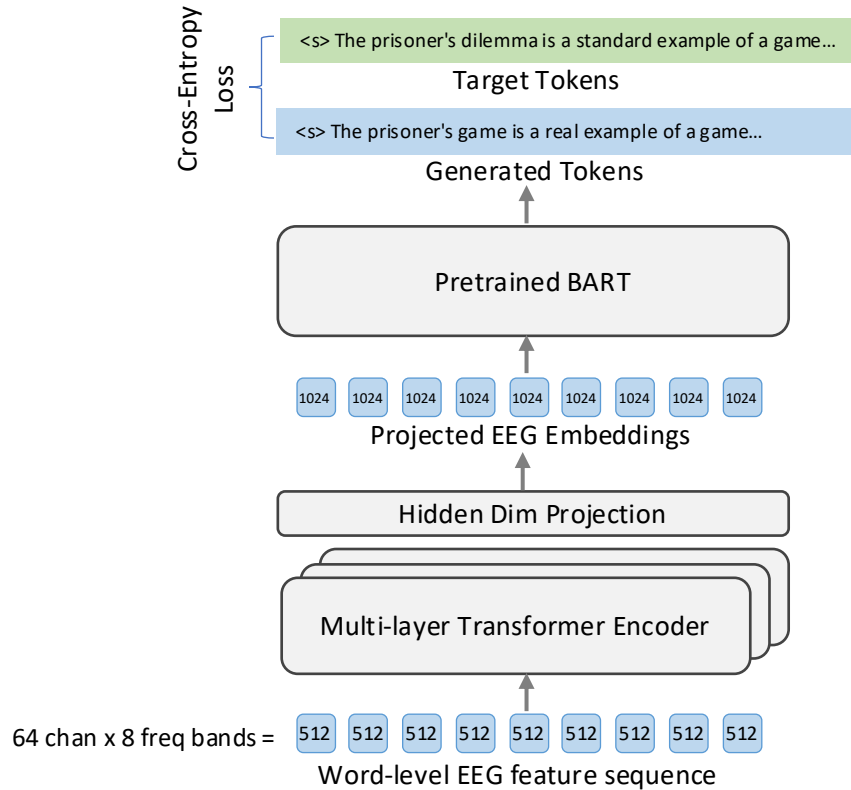
## Most Important Features

PC	Importance	Top Features
PC33	0.0148	fix_count, sacc_count, fix_R_duration, F6_theta1
PC34	0.0127	fix_count, sacc_count, fix_R_duration, CP5_theta2
PC61	0.0125	T8_theta1, FC6_theta1, PO8_beta2, CP3_theta1
PC31	0.0125	fix_count, sacc_count, fix_R_duration, F6_theta1
PC23	0.0115	Pz_theta1, P2_theta1, PO4_theta1, T7_alpha1

## Best-Performing Single-Subject Models



# EEG-to-Text Decoding



adapted from (Wang, 2022)  
<https://arxiv.org/abs/2112.02690>

Including MW Samples Degrades Performance

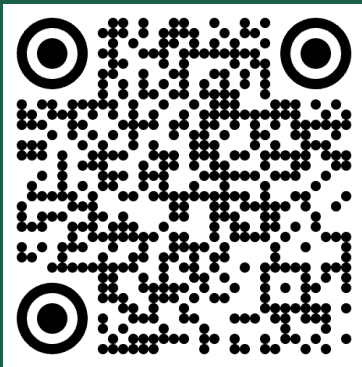
DATASET	# TRAIN	BLEU-N		ROUGE-1 F1
		N=1	N=2	
MW	6063	0.057	0.009	0.094
ALL	6063	0.168	0.054	0.241
NON-MW	6063	0.189	0.071	0.258
ZUCo	10710	0.401	0.231	0.301

# Thank You!

- Haorui Sun
- Ardyn Olszko
- Niharika Singh
- Jake Ruud
- Hazen Kellner
- Lincoln Lewis Esquerre
- **The rest of the Glass Brain Lab →**

Access ROAMM:

<https://openneuro.org/datasets/ds007629>



The authors acknowledge the Vermont Advanced Computing Center (VACC) at the University of Vermont for providing computational resources that have contributed to the research results reported within this paper.