Supporting Information for

Evaluation of Extreme Weather Impacts on Utility-scale Photovoltaic Plant Performance in the United States

Nicole D. Jackson and Thushara Gunda†
†Corresponding author email: tgunda@sandia.gov

This PDF file includes:

- Figures A1 to A11
- Tables A1 to A10

1 Supplemental Figures

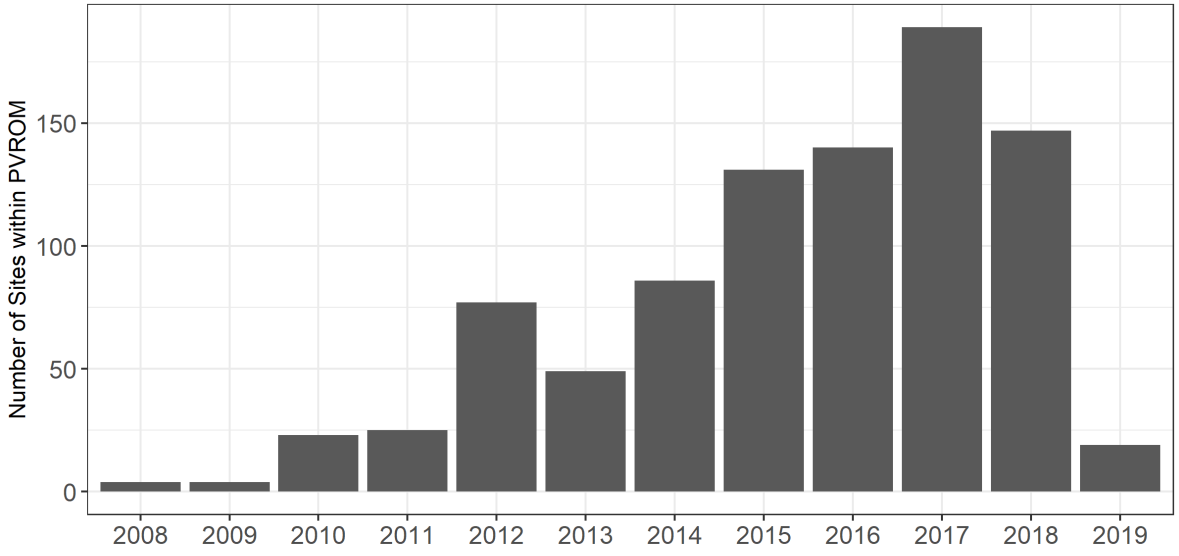

Figure A1: Histogram of commissioning dates of operations for sites within PVROM.

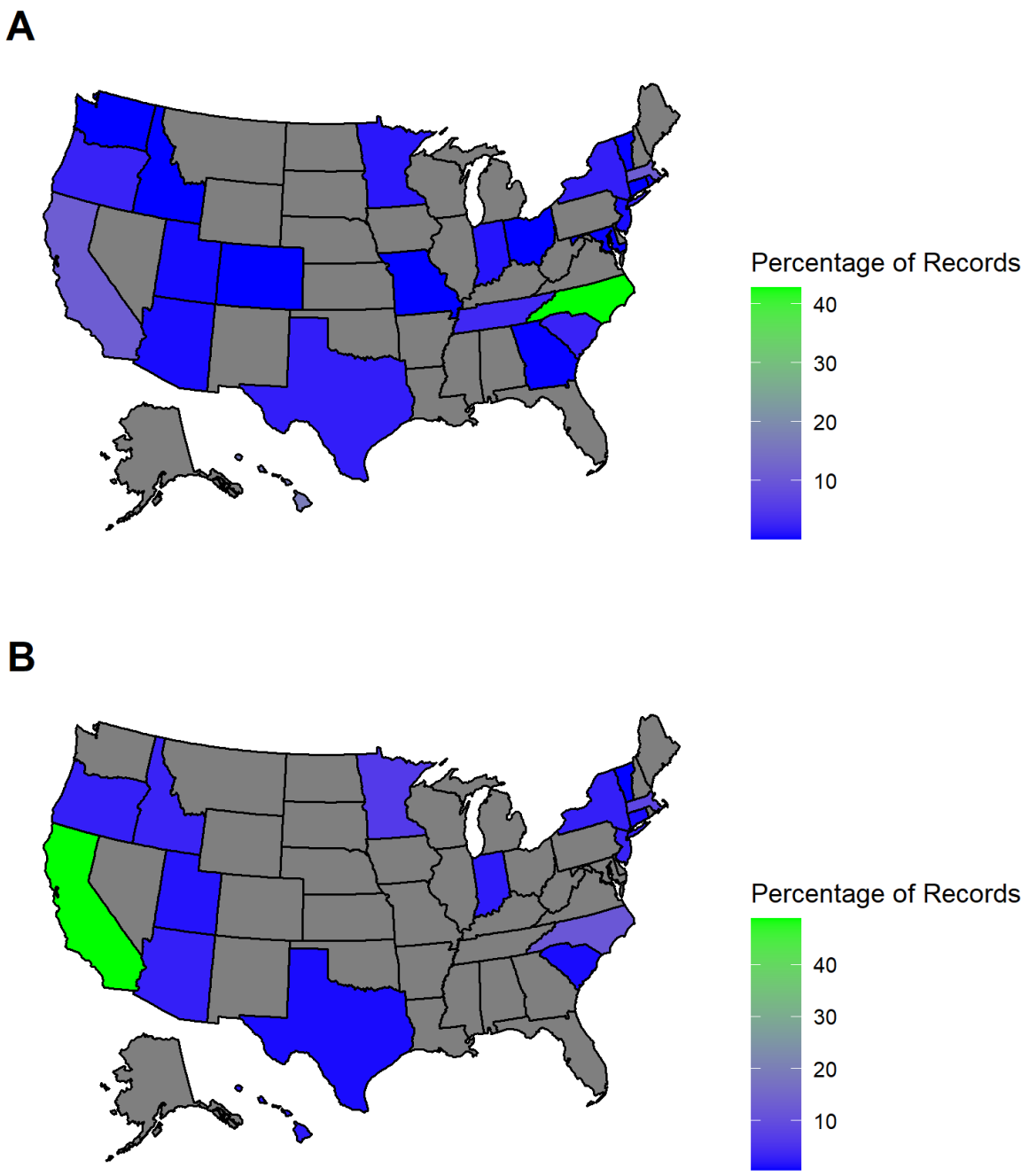

Figure A2: Geographic distribution of sites within PVROM analyzed for event identification using O&M records (A) and production analysis using performance and climate data (B).

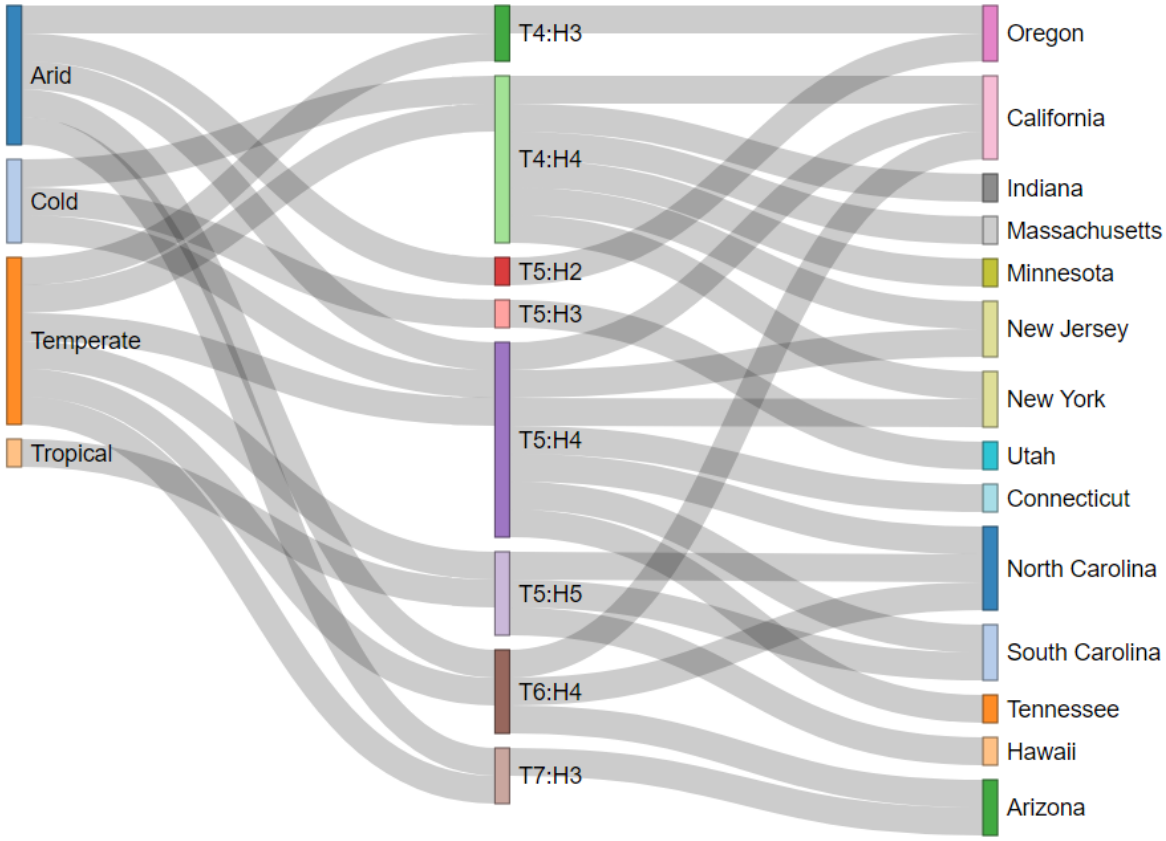

Figure A3: Distribution of sites with between Köppen Geiger regions (left; Peel et al., 2017), PV climate zones (middle; Karin et al., 2019), and the states where they are located.

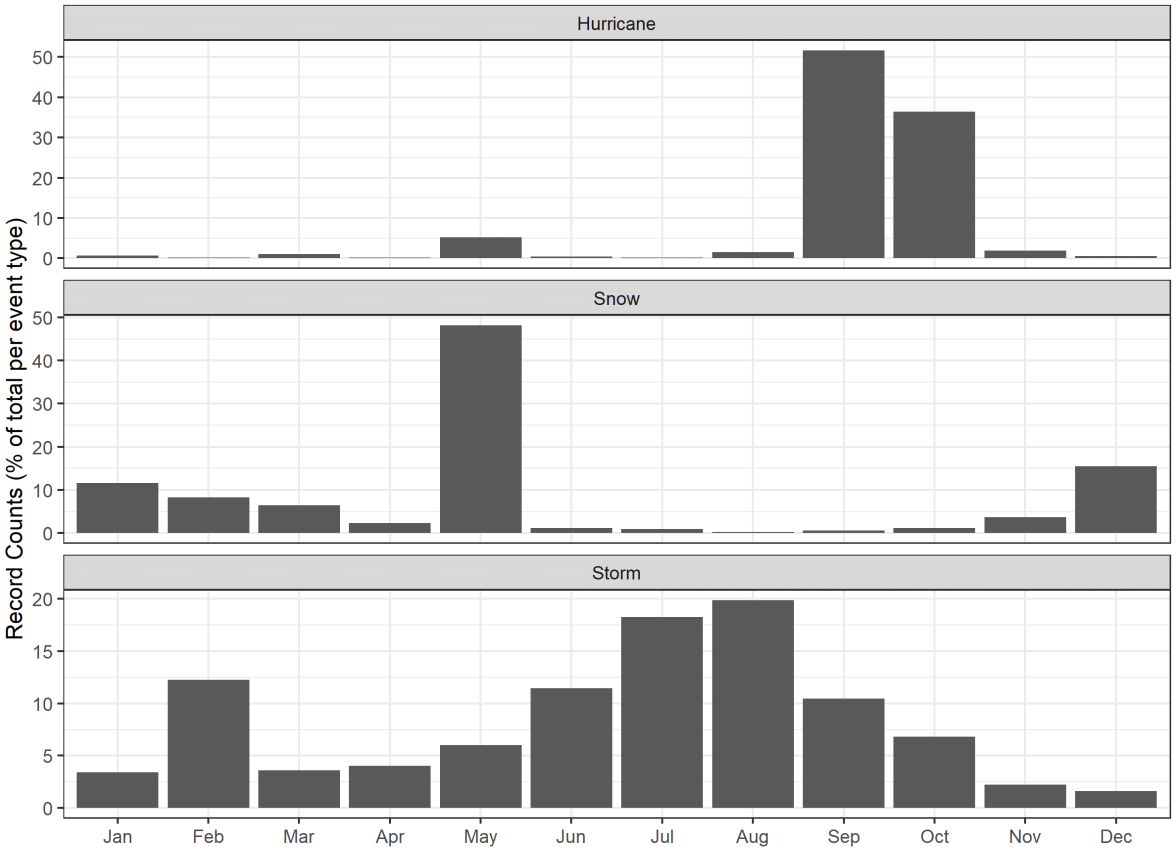

Figure A4: Distribution of events by month.

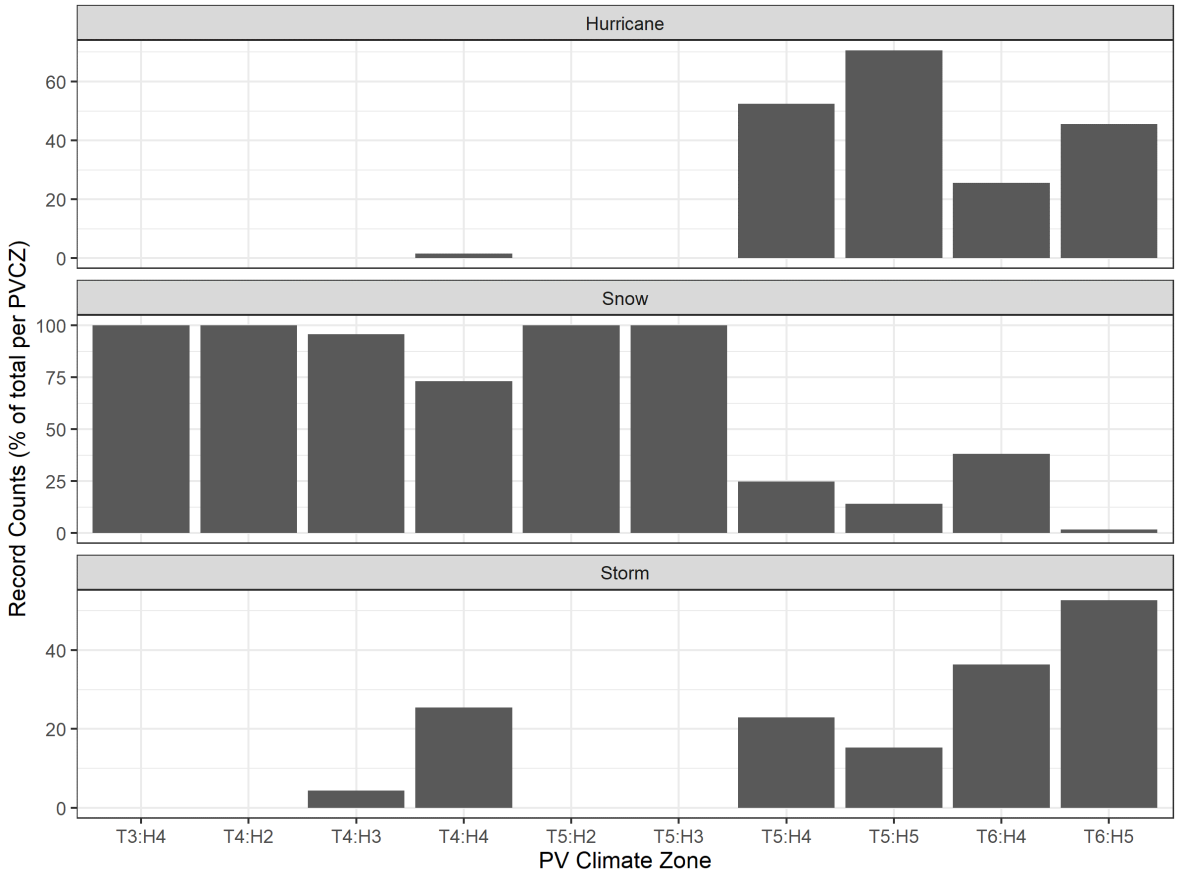

Figure A5: Summary of O&M Records by climate zone

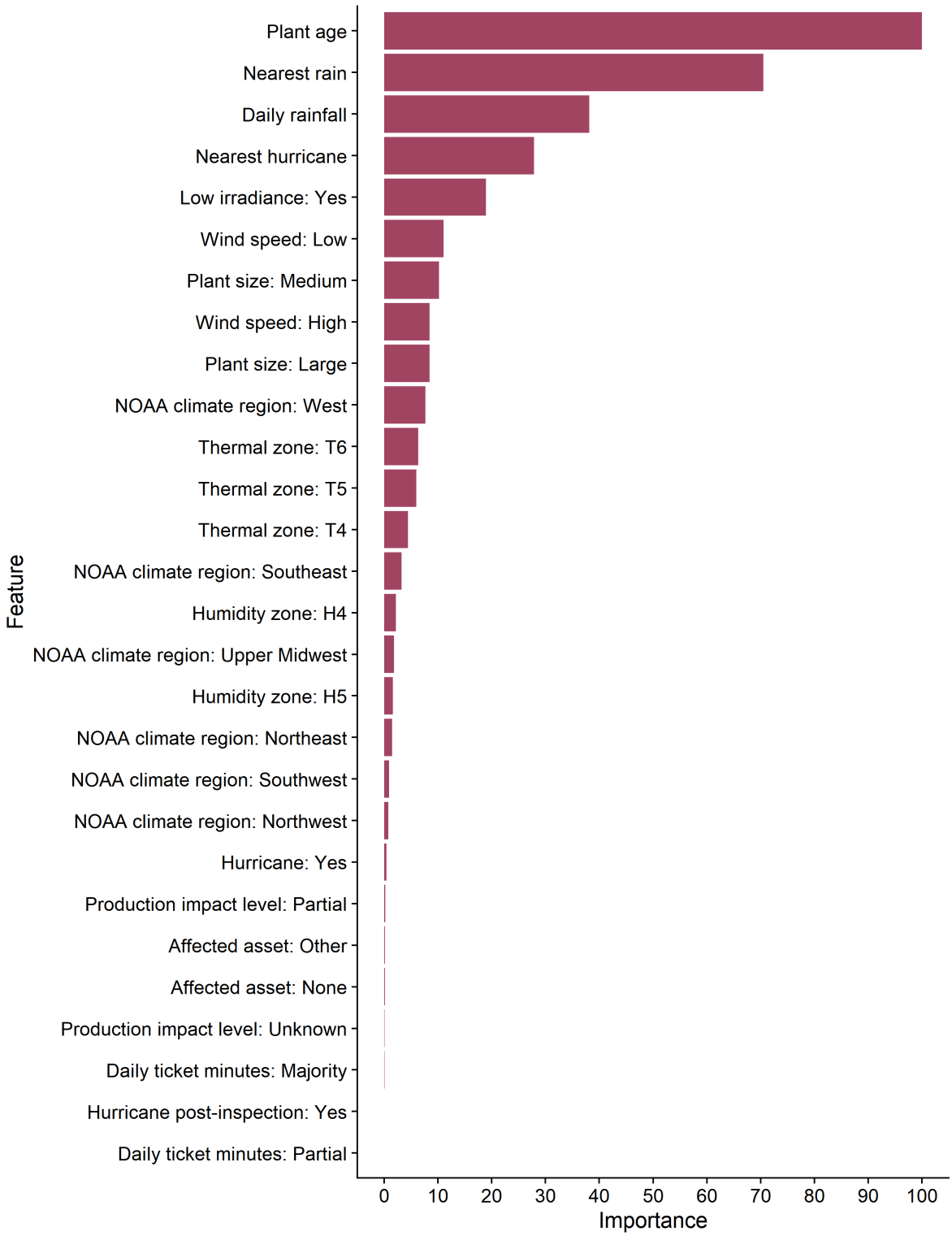

Figure A6: Key features of a random forest-based model hurricane-related variables. Features are presented in descending order of importance.

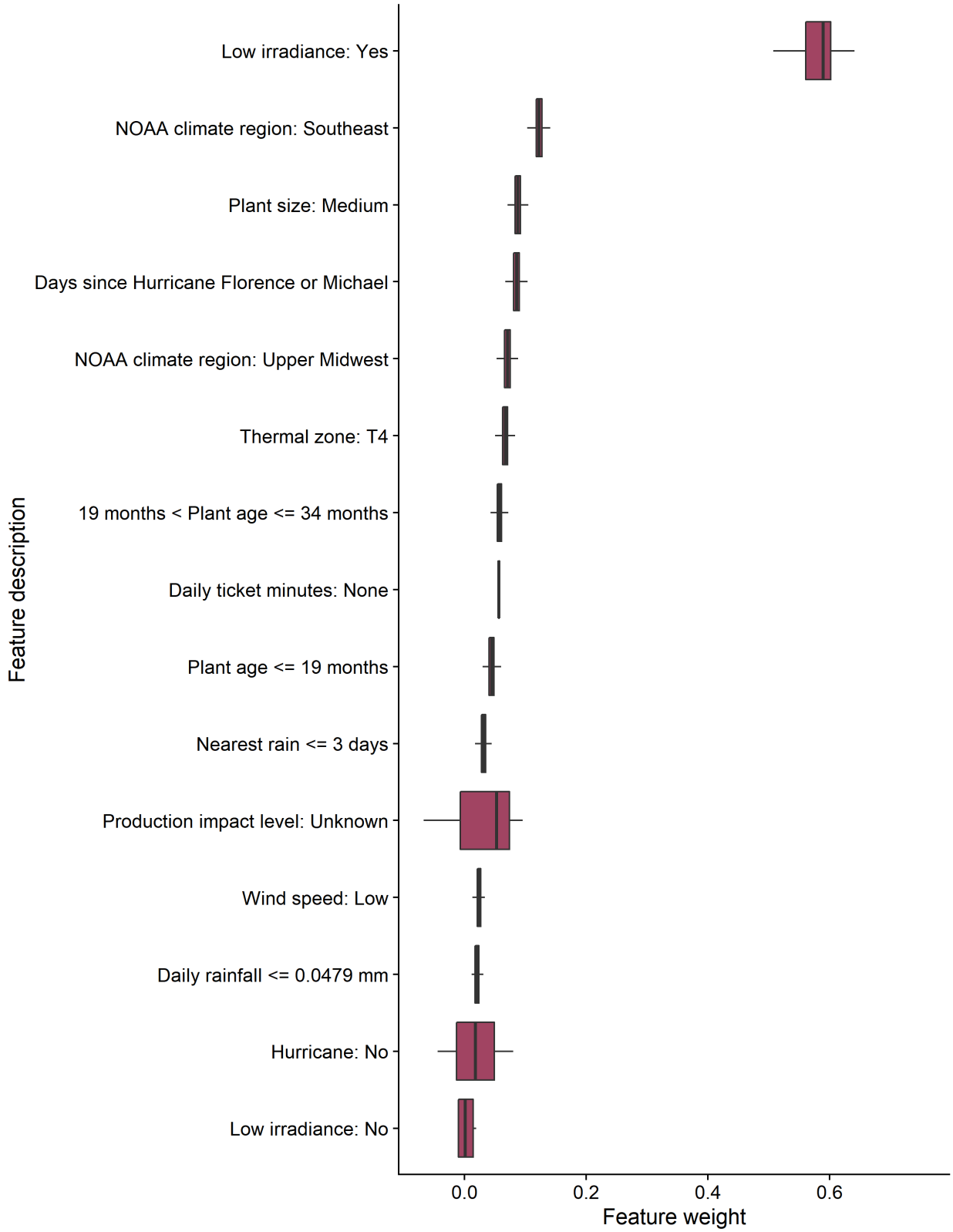

Figure A7: LIME explanations for low performance in the random forest-based models for hurricane events. Numerical thresholds are generated by LIME based on the data used in the analysis.

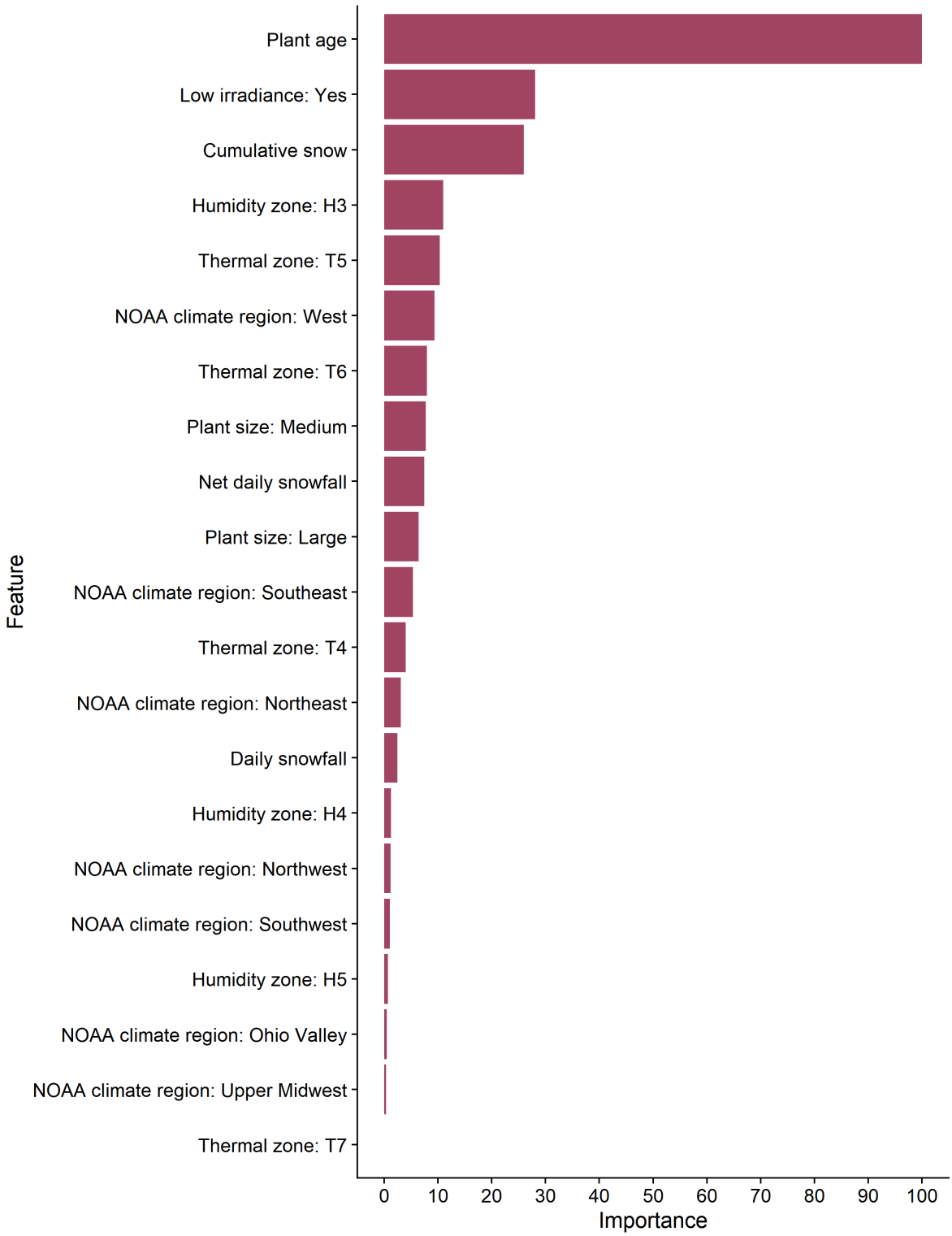

Figure A8: Key features of a random forest-based model snow-related variables. Features are presented in descending order of importance.

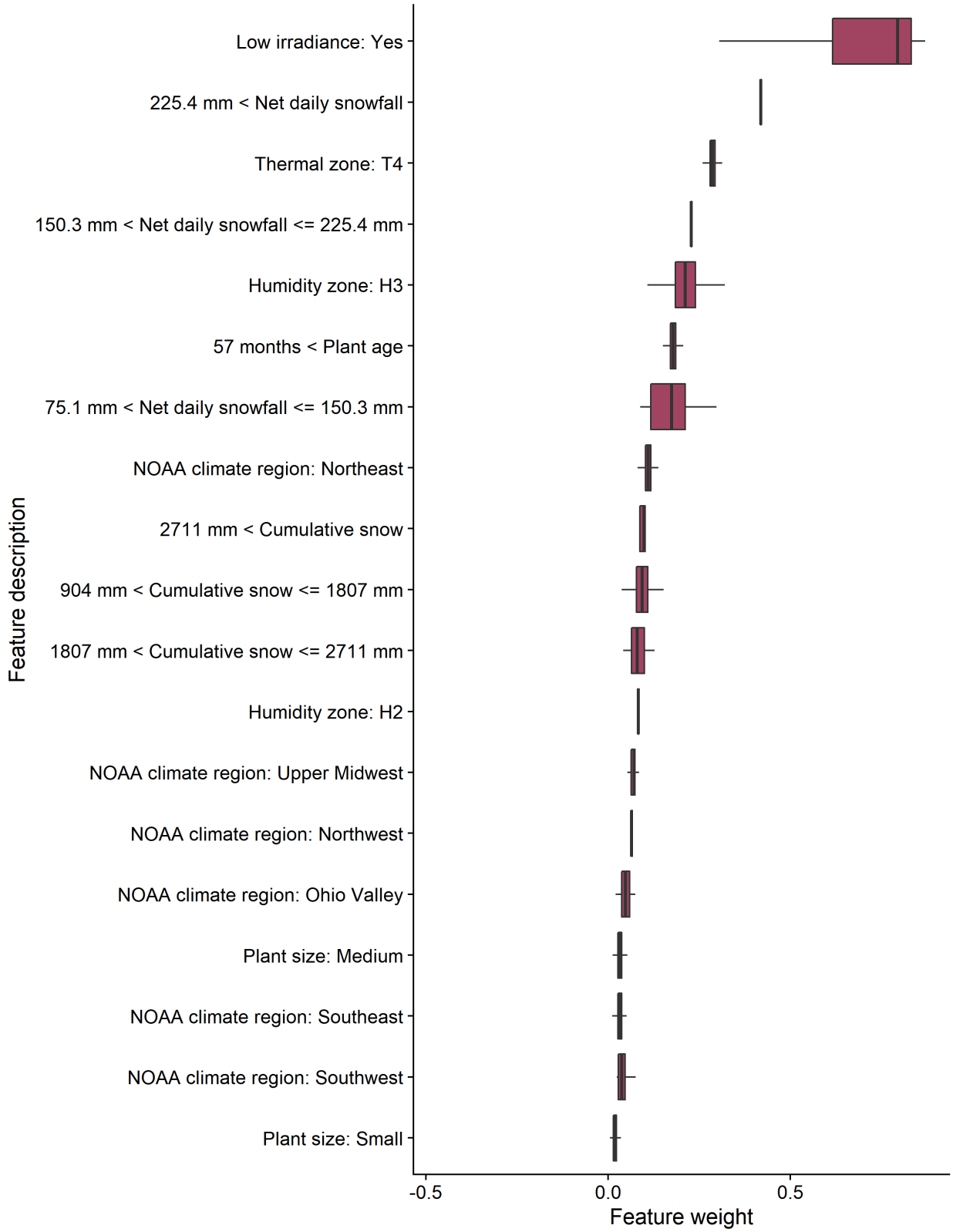

Figure A9: LIME explanations for low performance in the random forest-based models for snow events. Numerical thresholds are generated by LIME based on the data used in the analysis.

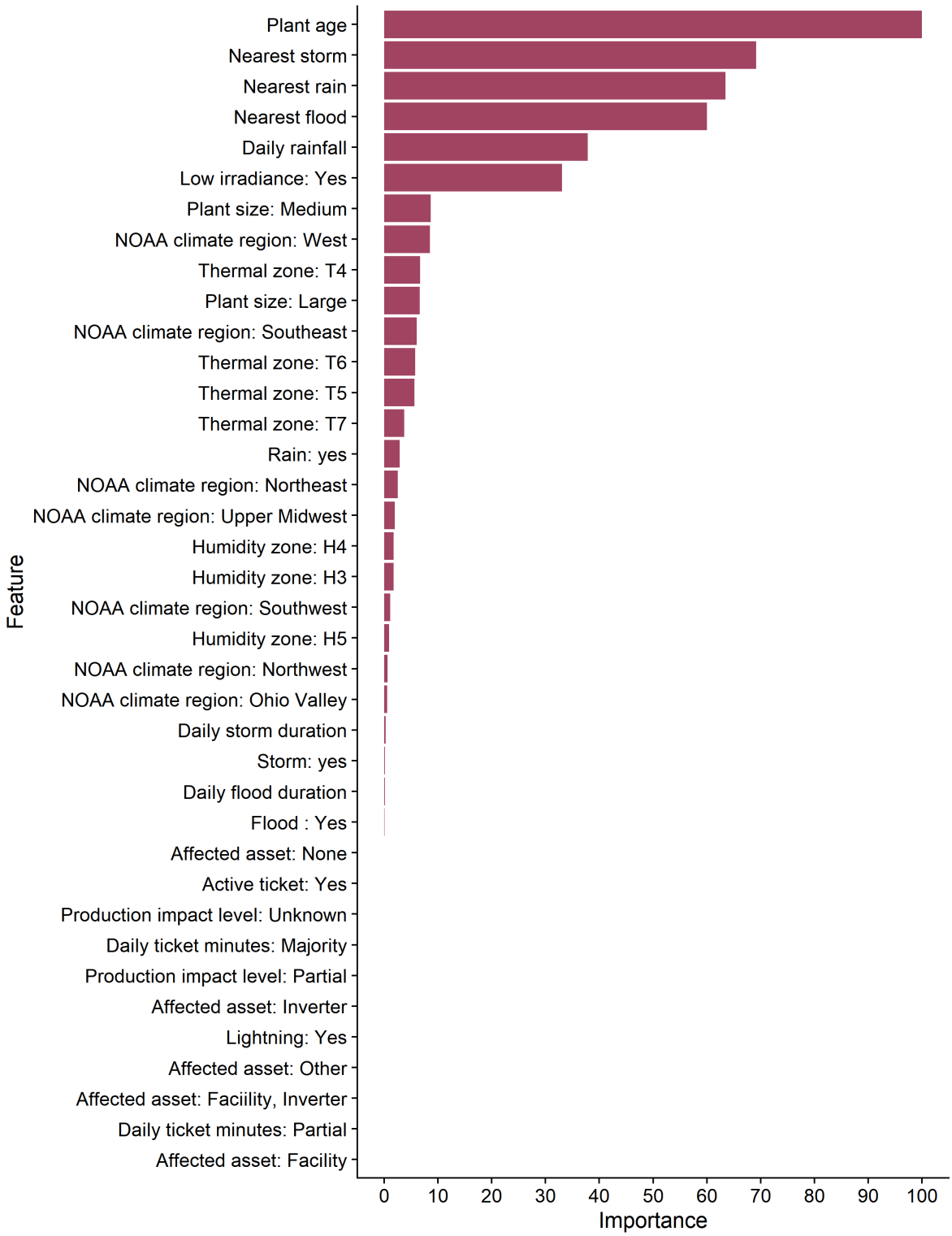

Figure A10: Key features of a random forest-based model storm-related variables. Features are presented in descending order of importance.

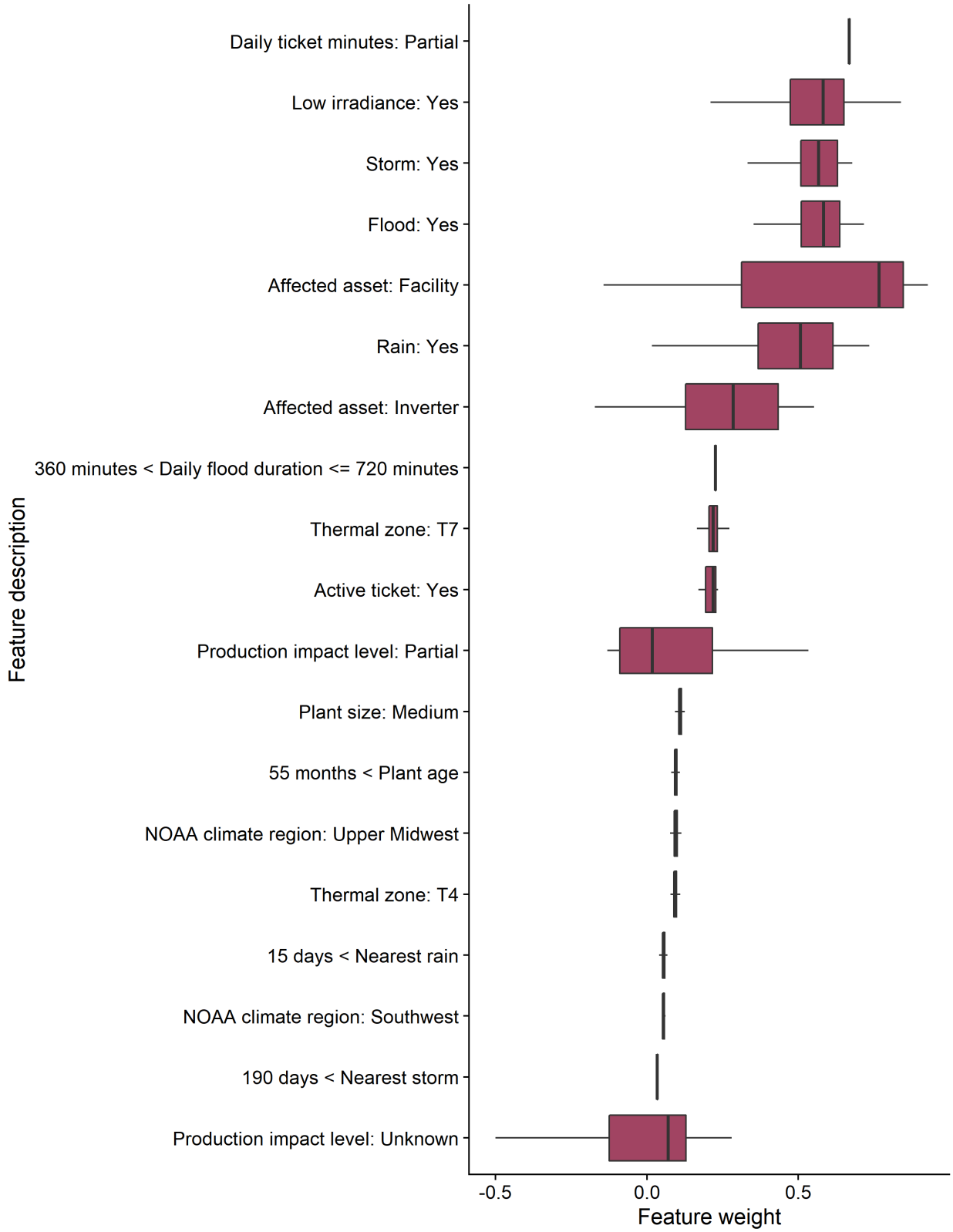

Figure A11: LIME explanations for low performance in the random forest-based models for storm events. Numerical thresholds are generated by LIME based on the data used in the analysis.

2 Supplemental Tables

Table A1: Terms Used for Filtering Relevant O&M Records. Initial lists were compiled from open source dictionaries (e.g., https://relatedwords.org). The regular expressions matching in R automatically captured plural variations of terms.

Parent category	Key terms	Notes
Extreme Weather	fire, flood, hail, hurricane, lightning, snow, storm, tropical	Words such as "cell" and "eye", which refer to specific parts of storm of hurricane were excluded since they are not often used in colloquial language.
Ambient Conditions	cloud, cold, conditions, cool, corrosion, erosion, heat, hot, moisture, rain, salt, spray, sun, temperature, temps, water, weather, wind	Intended to capture general conditions that might correlate to extreme weather events.

Table A2: Example records for event types. * indicates parent categories that were retained for further analysis.

Event Type	Asset	Description
Fire	Facility	"Ground fault caused fire in cable tray, destroying connections and cables."
Flood	Other	"Substation flooded due to storm."
Hurricane*	Facility	"Hurricane damage inspection. sWalked site looking for damage caused by hurricane."
Lightning	Inverter	"Inverter 4 experienced a ground fault on 4/24/19 ... Lightning Storms in area at time of Ground Fault. O&M Reset Inverter and Restarted with no issues."
Storm*	Facility	"Site not producing. Possible outage due to storm. Grid is OK.. . Targets cleared and site re-energized remotely."
Snow*	Facility	"[Site] producing near-zero due to recent snowfall. Ambient and module temperature are at -2 degrees Celsius. The site is still connected to the grid and communicating normally"
Wind	Module	"5 down series strings total found offline, 2 were repaired. The other three are offline dues to having one or more damaged modules on. One of them is in a location where the racking has become bent, due to hurricane-force winds."

Table A3: Summary date ranges for hurricanes overlapping with the PVROM database.

Hurricane	States affected	Event dates
Dorian	Florida, Georgia, North Carolina, South Carolina, Virginia	08/28/2019 – 09/10/2019
Florence	Georgia, Maryland, North Carolina, South Carolina, Virginia	09/07/2018 – 09/18/2018
Irma	Delaware, Florida, Georgia, Maine, Maryland, Massachusetts, North Carolina, South Carolina, Tennessee, Virginia	09/04/2017 – 09/17/2017
Michael	Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia	10/08/2018 – 10/16/2018

Table A4: Variables evaluated in the study. Type refers to if the variable is continuous (C), discrete/binned (D), or binary (B).

	Variable		
Company	Name	Type	Description
Climate	Irradiance	C, D	Total irradiance at the site for the day. Categories: low, medium, and high based on site
	Low irradiance	B	Indicator if Irradiance is categorized as "low"
	Snow	B	Indicator if there is snowfall for the given day
	Daily snowfall	C	Total daily snowfall at the site based on GHCN data
	Net daily snowfall	C	Difference snowfall and the amount of snow that would have melted
	Cumulative snow	C, D	Cumulative total snowfall since the start of the available production data.
	Hurricane	B	Indicator if there is a hurricane on the given day
	Nearest hurricane	C	Number of days since last recorded hurricane event. If no events were recorded, the last recorded event is set to the site's commissioning day.
	Wind speed	C, D	Average daily wind speed at site. Categories: low, medium, high
	Daily rainfall	C	Total daily rainfall at the site based on PRISM.
	Nearest rain	C	Number of days since last recorded rainfall. If no events were recorded, the last recorded event is set to the site's commissioning day.
	Storm	B	Indicator if there is a storm on the given day

Table A4 continued from previous page

	Variable		
Company	Name	Type	Description
	Daily storm duration	C	Total minutes across all storm events at site for a given day.
	Nearest storm	C	Number of days since last recorded storm event. If no events were recorded, the last recorded event is set to the site's commissioning day.
	Flood	B	Indicator if there is a flood on the given day
	Daily flood duration	C, D	Total minutes across all flood events at site for the given day
	Nearest flood	C	Number of days since last recorded flood event. If no events were recorded, the last recorded event is set to the site's commissioning day.
Operator/O&M	Company	D	Alpha-numeric company ID Categories: C1, C2, C3, C4, C5, C6
	Asset	D	Type of asset affected as identified by the O&M logs Categories: Combiner; Combiner, Facility; Combiner, Facility, Module; Combiner, Facility, Tracker; Combiner, Inverter; Combiner, Module; Facility; Facility, Inverter; Facility, Inverter, Tracker; Facility, Module; Facility, Tracker; Inverter; Inverter, Module; Inverter, Tracker; Module; Module, Tracker; Other; Tracker; Transformer
	Number of active tickets	B, C	Total number of active O&M tickets at the site for the given day Indicator if there any active O&M tickets at site for the given day
	Daily ticket duration	C, D	Total minutes across all active O&M tickets at site for the given day Categories: None (Daily ticket duration = 0), Partial (Daily ticket duration <720), Majority (720 >= Daily ticket duration <1440), Full (Daily ticket duration = 1440)
	Production impact level	D	Estimated production impact based on O&M records Categories: Full, Partial, Unknown, N/A
	Pre- Inspection	B	Indicator if inspection took place prior to hurricane event according to O&M logs

Table A4 continued from previous page

Company	Variable		
	Name	Type	Description
	Post-Inspection	B	Indicator if inspection took place after hurricane event according to O&M logs
	Lightning	B	Indicator if lightning was mentioned in O&M logs
Metadata	NOAA climate region	D	Climate classification system for the United States Categories: Hawaii, Northeast, Northwest, Ohio Valley, South, Southeast, Southwest, Upper Midwest, West
	PV Climate Zone	D	Climate zone based on a combination of temperature and humidity zones. See Karin et al. (2019) for details Categories: T1:H2, T4:H3, T4:H4, T5:H2, T5:H3, T5:H4, T5:H5, T6:H4, T7:H3
	Koppen-Geiger climate classification	D	System based on temperature and precipitation patterns. Categories: Continental, Dry, Temperate, Tropical
	State	D	State where site is located
	Plant size	D	DC nameplate power Categories: small, medium, large
	Plant age	C, D	Number of months since commissioning. Categories: Early (Age <12 months), Mid (12 months <= Age <60 months) , Late (>60 months)
	Array type	D	Type of array at site Categories: Fixed, Tracker, Mixed (combination of fixed and tracker), and NA

Table A5: Distribution of Weather Records. *Note: Indicates events retained for production impact analysis.

Extreme event	Number of tickets	Number of sites	Number of states
Fire	104	53	11
Flood	75	45	7
Hurricane*	916	297	6
Lightning	226	80	11
Storm*	484	192	13
Snow*	765	234	15
Wind	206	102	12

Table A6: Co-Occurrence of Weather Terms in O&M Records. For the statistical and ML analyses, storm-related assessments only contains tickets that did not include the term 'hurricane'.

Parent category	Terms	Percentage of event category
Hurricane	Hurricane + Storm	55%
Hurricane	Hurricane + Storm + Flood	4%
Hurricane	Hurricane + Wind	3%
Hurricane	Hurricane + Flood	0.7%
Hurricane	Hurricane + Storm + Lightning	0.4%
Storm	Storm + Lightning	21%
Storm	Storm + Flood	2%
Storm	Storm + Wind	5%
Snow	Snow + Storm	9%

Table A7: Spearman rank-order correlation results for each compound, extreme weather event analyzed. These values were used for relative comparisons to validate features identified through machine learning.

		r	p-value
Snow	Daily snowfall	-0.18	0
	Net daily snowfall	-0.17	0
	Number of snow tickets	-0.04	0.32
	Daily snow ticket duration	0.09	0.01
Hurricane	Nearest hurricane	-0.04	0
	Wind speed	0.10	0
	Number of hurricane tickets	-0.39	0
	Daily hurricane ticket duration	0.02	0.64
	Daily rainfall	-0.07	0
Storm	Nearest rain	0.00	0.66
	Nearest storm	-0.08	0
	Nearest flood	-0.10	0
	Daily storm duration	-0.22	0
	Daily flood duration	-0.24	0.03
	Number of storm tickets	-0.68	0
	Daily storm ticket duration	-0.15	0

Table A8: Confusion matrix of final selected random forest model for each compound, extreme weather event analyzed.

Term	Class	Snow	Hurricane	Storm
accuracy		0.64	0.69	0.73
kappa		0.46	0.54	0.60
sensitivity	low	0.62	0.68	0.75
specificity	low	0.86	0.86	0.88
pos_pred_value	low	0.68	0.70	0.76
neg_pred_value	low	0.82	0.84	0.88
precision	low	0.68	0.70	0.76
recall	low	0.62	0.68	0.75
f1	low	0.65	0.69	0.76
prevalence	low	0.33	0.33	0.33
detection_rate	low	0.20	0.23	0.25
detection_prevalence	low	0.30	0.32	0.33
balanced_accuracy	low	0.74	0.77	0.82
sensitivity	medium	0.65	0.69	0.73
specificity	medium	0.80	0.83	0.86
pos_pred_value	medium	0.62	0.67	0.72
neg_pred_value	medium	0.82	0.84	0.86
precision	medium	0.62	0.67	0.72
recall	medium	0.65	0.69	0.73
f1	medium	0.64	0.68	0.72
prevalence	medium	0.33	0.33	0.33
detection_rate	medium	0.22	0.23	0.24
detection_prevalence	medium	0.35	0.35	0.34
balanced_accuracy	medium	0.73	0.76	0.79
sensitivity	high	0.66	0.70	0.72
specificity	high	0.81	0.86	0.86
pos_pred_value	high	0.63	0.71	0.72
neg_pred_value	high	0.83	0.85	0.86
precision	high	0.63	0.71	0.72
recall	high	0.66	0.70	0.72
f1	high	0.65	0.71	0.72
prevalence	high	0.33	0.33	0.33
detection_rate	high	0.22	0.24	0.24
detection_prevalence	high	0.35	0.33	0.33
balanced_accuracy	high	0.73	0.78	0.79

Table A9: Summary of feature importance weights for each compound, extreme weather event analyzed. Features showing zero are only zero due to rounding.

Feature	Snow	Hurricane	Storm
Active ticket: Yes	0.02		3.72
Affected asset: Facility	0.4	1.84	2
Affected asset: Faciility, Inverter	0.53		0.57
Affected asset: Inverter	0.71	1.62	0.96
Affected asset: None	1.1	0.93	1.18
Affected asset: Other	1.23	0.79	0.67
Plant size: Large	1.27	2.21	1.79
Plant size: Medium	2.48		
Daily ticket minutes: Majority	4	4.47	6.71
Daily ticket minutes: Partial	3.09	1.47	2.56
Wind speed: High	5.35	3.28	6.05
Wind speed: Low	6.4	8.44	6.62
Cumulative snow	9.36	7.66	8.51
Daily flood duration	7.5		
Daily storm duration	7.99	6.34	5.8
Flood : Yes	7.75	10.23	8.69
Humidity zone: H3	10.96		1.77
Humidity zone: H4	10.39	6.01	5.64
Humidity zone: H5	26		
Hurricane: Yes	28.09	18.94	33.12
Hurricane post-inspection: Yes	100	100	100
Lightning: Yes		0	0
Low irradiance: Yes		0.01	
Nearest flood		0.11	0.03
Nearest hurricane		0.11	0.03
Nearest rain		0.15	0.01
Nearest storm		0.13	0.03
NOAA climate region: Northeast		0.22	0.02
NOAA climate region: Northwest		0.43	
NOAA climate region: Ohio Valley		8.45	
NOAA climate region: Southeast		11.06	
NOAA climate region: Southwest		27.91	
NOAA climate region: Upper Midwest		38.13	37.84
NOAA climate region: West		70.51	63.5
Plant age			0.01
Production impact level: Partial			0
Production impact level: Unknown			0.01
Rain: yes			0.02
Daily rainfall			0.03
Daily snowfall			0.11
Storm: yes			0.15
Thermal zone: T4			0.2
Thermal zone: T5			0.32
Thermal zone: T6			2.89
Thermal zone: T7			60.03
Net daily snowfall			69.19

Table A10: Summary of feature influence and mean weight for each compound, extreme weather event analyzed.

Feature description	Snow	Hurricane	Storm
Active ticket: Yes			
Affected asset: Facility			
Affected asset: Inverter			
Affected asset: Other			
Plant size: Medium	0.033	0.088	0.033
Plant size: Small	0.018		0.018
Daily ticket minutes: Majority			
Daily ticket minutes: None		0.057	
Daily ticket minutes: Partial			
Wind speed: Low		0.024	
1807 mm <Cumulative snow <= 2711 mm	0.084		0.084
2711 mm <Cumulative snow	0.095		0.095
904 mm <Cumulative snow <= 1807 mm	0.095		0.095
360 minutes <Daily flood duration <= 720 minutes			
Daily storm duration <= 360 minutes			
Flood: Yes			
Flood: No			
Humidity zone: H2	0.083		0.083
Humidity zone: H3	0.211		0.211
Hurricane: No		0.019	
Lightning: No			
Lightning: Yes			
Low irradiance: Yes	0.706	0.596	0.706
Low irradiance: No		0.003	
647 days <Nearest flood			
Days since Hurricane Florence or Michael		0.086	
15 days <Nearest rain			
18 days <Nearest rain			
Nearest rain <= 3 days		0.032	
NOAA climate region: Northeast	0.11		0.11
NOAA climate region: Northwest	0.064		0.064
NOAA climate region: Ohio Valley	0.043		0.043
NOAA climate region: Southeast	0.032	0.123	0.032
NOAA climate region: Southwest	0.024		0.024
NOAA climate region: Upper Midwest	0.069	0.071	0.069
19 months <Plant age <= 34 months		0.058	
28 months <Plant age <= 45 months			
55 months <Plant age			
57 months <Plant age	0.178		0.178
Plant age <= 19 months		0.045	
Plant age <= 28 months			
Production impact <= Full			
Production impact <= Unknown		0.027	
Rain: Yes			
Daily rainfall <= 0.0479 mm		0.021	
Daily rainfall <= 0.0784 mm			
Daily snowfall <= 68.9 mm			
Storm: Yes			
Thermal zone: T1			

Table A10 continued from previous page

Feature description	Snow	Hurricane	Storm
Thermal zone: T4	0.286	0.067	0.286
Thermal zone: T7			
150.3 mm <Net daily snowfall <= 225.4 mm	0.228		0.228
225.4 mm <Net daily snowfall	0.419		0.419
75.1 mm <Net daily snowfall <= 150.3 mm	0.176		0.176
75.5 mm <Net daily snowfall <= 151.0 mm			