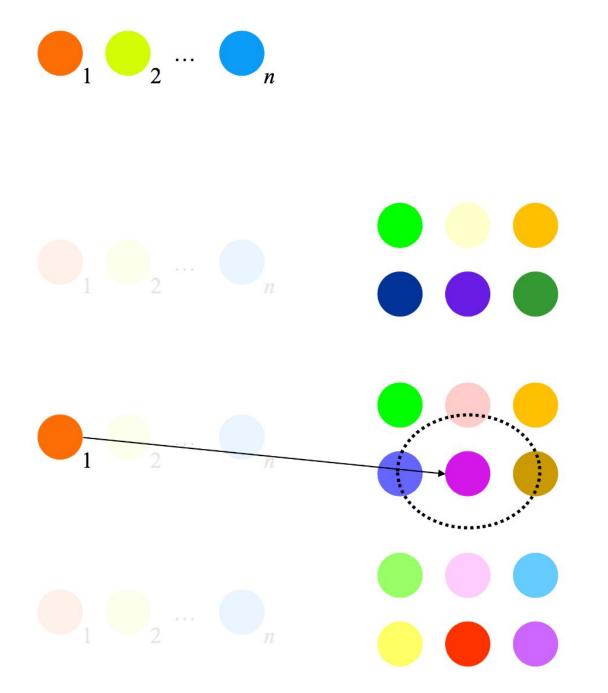
Investigating precipitation regimes using self-organizing maps

Why Is It Important?

Heavy precipitation can have major impacts on infrastructure, agriculture, and society, and different regions have different criteria for heavy precipitation. This therefore necessitates accurate background knowledge and forecasting of the weather regimes that can cause heavy precipitation for a given region, as well as common regimes that precede and follow them that can aid in forecasting. However, precipitation and other meteorological datasets can be large, making identification of important features difficult without the use of machine learning. In particular, self-organizing maps (SOMs) can train on large datasets and parse out the most common and most distinct patterns (Figure 1).

This investigation of precipitation regimes involves training SOMs on Multi-Radar Multi-Sensor (MRMS) observed precipitation data from January 2021 to December 2023. Six SOMs are trained on MRMS 1-h and 24-h quantitative precipitation estimate (QPE) data, with each SOM focused on one of six regions within CONUS to parse out storm types affecting each region. SOMs are able to cluster regimes with heavier precipitation, allowing investigation of where the heaviest precipitation falls for each region on average, when it is most likely to occur, how long it typically persists before transitioning to another regime, and regime-based numerical weather prediction (NWP) model forecast errors.



Step 0: Dataset consists of *n* examples of a color object, described by a 3×1 Red-Green-Blue vector (e.g. [0.5;1;1]). For a more complex example using a weather map, the object would be $(w^*h^*m) \times 1$, where w is the width of the field, h is the height of the field, and *m* is the number of variables.

Step 1: Initialize 3×2 SOM grid, either randomly or using the principal components of the original dataset. The size of the grid is decided at the beginning. There are now 6 nodes, each 3×1 in size to match the dimensions of the input data.

Step 2: Loop through each of the *n* objects in the input dataset. For each, determine which SOM node is the best matching unit (node in center of black dashed oval). Update the weights of that node and its neighbors (any nodes within the oval) so those nodes more closely resemble the input object.

Step 3: Continue looping through the dataset until the max iterations *p* have been reached, and the SOM nodes summarize the original dataset well.

Figure 1. A diagram of how SOMs work using the example of identifying the most representative colors from a dataset of RGB values. (From Naegele et al. 2024)

Data and Methods

- Multi-Radar Multi-Sensor (MRMS; Zhang et al. 2016)
 - 1-hr & 24-h Quantitative Precipitation Estimate (QPE)
- 6 sub-CONUS domains (here the focus is on South-Central) Jan 2021 - Dec 2023
- Data grid thinned by retaining every 10th grid point (10-km resolution) during training to save computational cost
- MiniSom (Vettigli 2018)
 - Hyperparameters (Table 1) resulted in lowest quantization and topographic errors

Grid Size	Learning Rate	Sigma	Еро		
5×5	0.25	1.75	50		

Table 1. MiniSom Hyperparameters

References

Naegele, S. J. A. Lee, S. J. Greybush, G. S. Young, S. E. Haupt, 2024: Identifying wind regimes near Kuwait using self-organizing maps. J. Renewable Sustainable Energy, **16**, 026501, https://doi.org/10.1063/5.0152718.

Vettigli, G., 2018: MiniSom: minimalistic and NumPy-based implementation of the Self Organizing Map. Available at

https://github.com/JustGlowing/minisom/, accessed 03 Apr 2024. Zhang, J., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. Bull. Amer. Meteor. Soc., 97, 621–638, https://doi.org/10.1175/BAMS-D-14-00174.1.



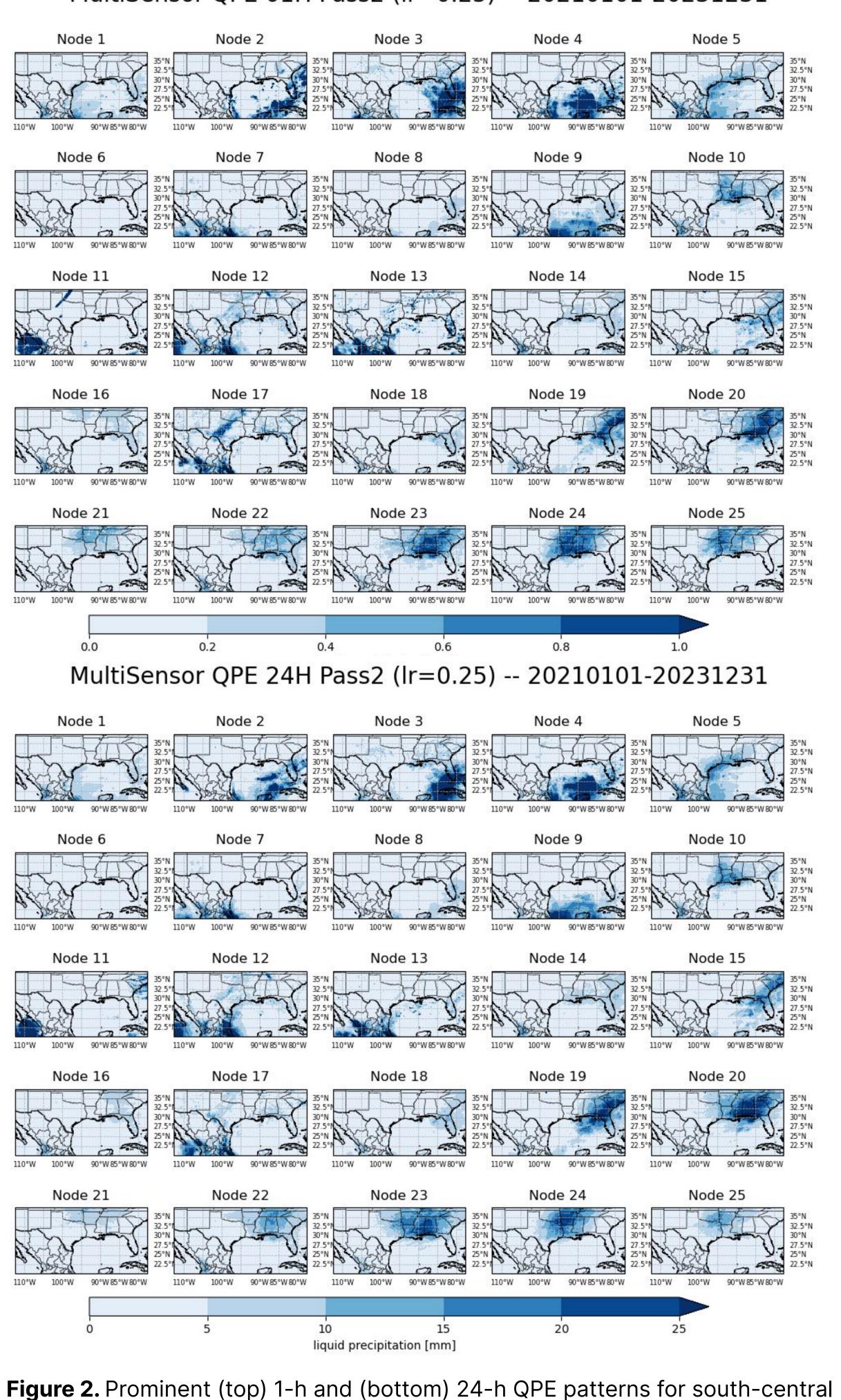
oochs

500

Spatial and Seasonal Patterns

- Several nodes represent heavier precipitation (Figure 2) • Nodes 2, 3, 4, 11, 12, and 13 indicate average 1-hr quantitative precipitation estimate (QPE) > 1 mm and average 24-hr QPE > 25 mm somewhere in eastern Gulf of Mexico, Florida, and Caribbean (nodes 2, 3, and 4), as well as central Mexico (nodes 11, 12, and 13)
- Nodes 9, 10, 15, 17, and 19–25 represent more moderate-heavy 1-h and 24-h QPE, primarily over southeastern states
- The remaining nodes represent drier regimes with lighter mean QPE and lower node-based standard deviations of QPE
- Nodes with greatest QPE are generally the least common, but they tend to occur more frequently in summer and autumn months (Figure 3) • The moderate-heavy QPE nodes are more common in spring-autumn
- Drier nodes can occur all year • Node 6 in particular has little to no QPE anywhere and is the most common node, occurring most often in winter months

MultiSensor QPE 01H Pass2 (lr=0.25) -- 20210101-20231231



CONUS from January 2021 to December 2023 after SOM training.



Ν		100000	1000000000000	A STATE OF A	CONTRACTOR CONTRACTOR	100000000	100000000000000000000000000000000000000	Accessed and an Accession	and the second se	00000000		100000000000000000000000000000000000000	=0.25))
1-	16	52	108	151	108	31	158	162	34	43	90	148	49	
2-	0	0	0	6	11	0	0	0	3	5	0	0	0	
3-	10	0	25	4	28	0	52	38	13	4	26	18	0	- 80
4 -	20	0	49	18	1	0	0	19	0	11	0	0	20	
5 -	71	185	397	195	112	9	148	171	63	137	126	169	11	-70
6-	496	683	351	704	875	879	315	602	750	536	365	452	180	//
7 -	0	25	8	18	0	6	38	140	0	0	35	78	25	
8-	44	149	23	205	251	113	40	221	192	111	84	268	94	- 60
9-	25	0	65	111	58	10	34	170	78	35	68	92	76	00
10 -	63	77	94	22	13	28	134	33	18	21	34	60	0	
2 11 -	0	0	0	1	0	0	0	0	0	0	0	8	0	- 50
12 -	30	0	32	22	19	11	16	89	0	21	24	79	0	
13-	12	0	0	0	0	2	0	14	0	0	24	11	0	
14	48	43	180	116	84	44	218	28	48	110	99	162	36	- 40
n 15 -	82	17	5	8	28	23	7	51	20	10	26	18	13	
16 -	109	132	104	164	93	166	483	60	162	273	431	137	14	
17 -	4	0	11	25	3	0	37	3	0	39	9	22	0	- 30
18 -	84	78	169	59	26	128	145	31	96	138	95	51	11	
19-	73	9	63	36	29	87	0	28	90	47	113	23	59	
20 -	47	60	108	40	64	168	0	38	72	42	17	20	53	- 20
21 -	51	136	26	60	97	152	86	49	149	180	182	158	9	
22 -	40	63	96	109	72	52	172	21	48	66	192	33	7	10
23 -	24	100	159	18	48	88	53	36	73	90	75	16	11	- 10
24 -	55	221	126	46	90	138	65	126	142	190	46	124	13	
25 -	12	173	2	43	46	70	4	47	106	99	44	35	5	0
$\frac{25 - 12}{v^{1/3}} \frac{173}{2} \frac{2}{43} \frac{43}{46} \frac{70}{70} \frac{4}{47} \frac{47}{106} \frac{106}{99} \frac{99}{44} \frac{44}{35} \frac{5}{5} - 0$														

Figure 3. Number of hours within a given season represented by each SOM node.

		Noc	de 3	trar	nsiti	on p	orob	abili	ty ['	%] v	vith	tran	sitic	on pe	eriod	d ler	ngth		
1.	0.0	0.0	0.0	0.0	0.5	0.9	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.8	11.5	8.7	0.0	
2-	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
3.	95.0	89.9	84.9	71.6	59.2	49.1	32.1	20.6	8.3	1.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	- 80
5	0.0	0.0	0.0	0.9	1.4	0.9	0.0	0.9	6.4	8.7	0.0	0.0	0.0	0.0	0.0	0.0	8.7	3.7	00
6	0.0	0.0	0.0	0.5	3.7	8.7	22.5	28.9	33.9	42.2	29.4	23.4	50.5	47.2	47.7	30.3	48.6	43.6	
7.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.8	0.0	0.0	0.0	0.0	0.0	0.0	1.4	12.4	9.2	2.3	
8		6.4	10.1	20.6	27.1	29.4	27.1	20.2	18.8	8.7	0.0	0.0	2.3	11.5	8.7	0.0	2.3	4.6	
e 9.	0.9	1.8	2.3	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.8	- 60
⊕ 10-	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.5	0.0	0.0	0.0	0.0	0.5	2.3	6.0	0.0	0.0	- 60
2 11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
Σ 12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	4.6	0.0	0.0	3.7	
O 13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.7	0.0	0.0	0.0	0.0	
ω 14	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.8	6.9	6.0	6.4	4.1	9.2	5.5	6.0	0.0	11.9	
n 15	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	1.4	7.8	2.8	0.0	0.0	0.5	0.0	0.9	0.0	- 40
Į 16	0.0	0.0	0.0	0.0	0.0	0.0	0.5	3.2	6.9	11.9	22.9	25.7	8.3	3.2	3.2	15.6	1.8	2.8	
1/	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	
18	0.0	0.0	0.0	0.0	0.0	0.0	0.9	3.2	5.5	6.0	7.8	6.4	5.0	1.8	0.0	0.5	5.5	6.9	
19	0.9	1.8	2.8	5.5	8.3	11.0	15.1	17.9	11.9	4.6	0.0	1.8	0.0	6.0	0.0	0.0	1.4	0.0	
20	2003 B	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.3	0.0	0.0	0.0	0.0	4.1	0.0	0.0	- 20
21	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.3	5.0	8.3	8.3	6.0	0.5	0.0	3.7	1.4	7.3	
22		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	8.7	11.0	8.7	1.4	0.0	2.8	4.1	
23		0.0	0.0	0.0	0.0	0.0	1.4	2.8	2.8	2.3	9.2	8.3	2.8	2.3	3.7	1.8	0.0	0.0	
24	and the second second	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	6.0	5.0	8.7	4.6	11.5	4.6	8.7	0.0	
25	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.2	1.4	0.0	1.8	3.7	0.0	0.9	- 0
	\mathbf{r}	r	З	6	9	~~	~~~	20	30	200	12	°6	220	24.4	16° .	Sr .	2° -	200	
									time	hours]								

Figure 4. Probability (in %) of node 3 persisting or transitioning to another node after a given number of hours.

Node Transitions

- followed by node 16 after 72–96 hours • Nodes 6, 8, and 16 represent nodes with little to no average

Future Work

- Continue investigating other sub-CONUS domains
- model performance

Acknowledgements

This research was supported in part by NOAA cooperative agreement NA22OAR4320151, CIESRDS (Cooperative Institute for Earth System Research and Data Science).



• Focusing on node 3 (heavier average precipitation in south Florida) and calculating the probability of transitioning to another node over time • Node 3 persists for about 6 hours before transitioning (Figure 4) • The most common nodes node 3 transitions to are nodes 6, 8, and 19,

precipitation, and thus a weakening of the storms present in node 3 • Node 19 indicates a northward progression of mean precipitation

• Filter data (ARI threshold) for more focus on regimes with precipitation

• Cluster NWP QPF using MRMS-trained SOM nodes and compare with MRMS to assess model forecast skill under difference precipitation regimes

• Add NWP model error as a training variable to further assess regime-based