

Table 3: Error metrics evaluated on Training, Validation, and Testing Dataset

<b>Models</b>	<b>Train:</b>	<b>Val:</b>	<b>Test:</b>
Normal Short	(185.30, 227.07, 5.78%)	(156.81, 199.68, 4.89%)	(270.41, 319.46, 8.71%)
<b>Bayesian Short</b>	(203.08, 247.24, 6.27%)	(184.82, 229.84, 5.79%)	(281.64, 334.32, 8.98%)
Normal Long	(165.35, 199.79, 5.04%)	(170.86, 202.79, 5.68%)	(165.31, 199.79, 5.04%)
<b>Bayesian Long</b>	(256.62, 286.64, 7.39%)	(167.82, 216.20, 5.56%)	(168.89, 200.05, 5.35%)

Note: The errors in the tuple are formatted as *MAE*, *RMSE*, *MAPE*

Table 4: Encapsulation Frequency with Confidence Interval Bounds

<b>Models</b>	<b>Train:</b>	<b>Val:</b>	<b>Test:</b>
Bayesian Short	(97.16%, 99.60%, 99.74%)	(96.89%, 99.75%, 99.93%)	(90.04%, 97.95%, 98.96%)
Bayesian Long	(98.5%, 100%, 100%)	(99.3%, 100%, 100%)	(99.19%, 100%, 100%)

Note: The encapsulation frequency is formatted as (70% CI, 90% CI, 95% CI)

Table 5: Exceedance Frequency from 70% Confidence Interval Bounds

<b>Models</b>	<b>Train:</b>	<b>Val:</b>	<b>Test:</b>
Bayesian Short	(2.51%, 0.32%)	(2.66%, 0.44%)	(8.04%, 1.92%)
Bayesian Long	(1.49%, 0.006%)	(0.60%, 0.07%)	(0.23%, 0.58%)

Note: The statistics is formatted as (upper percentage exceedance, lower percentage exceedance)

Table 6:  $R^2$  coefficient values and mean bias from Years: 2019 - 2021, Months: 3-5

<b>Models</b>	<b>2019:</b>	<b>2020:</b>	<b>2021:</b>
Normal Short	(0.51, -8.50)	(0.45, -35.53)	(0.52, -22.26)
<b>Bayesian Short</b>	(0.51, -73.54)	(0.38, -81.68)	(0.37, -58.55)
Normal Long	—	(0.04, 337.19)	(0.48, 274.04)
<b>Bayesian Long</b>	—	(0.79, 71.83)	(0.87, 21.48)