# 1 Appendix

Table 1: Comparison of Chronos models' performance using various metrics for different datasets.

| Dataset | sMAPE | NRMSE | MAE |
|---|---|---|---|
| **Commercial Buildings** | | | |
| BDG-2 | 9.95 | 15.72 | 586.39 |
| Buildings-900K-test | 17.20 | 27.96 | 625.50 |
| Electricity | 5.94 | 9.00 | 1031.85 |
| **Residential Buildings** | | | |
| LCL | 41.04 | 84.01 | 7.37 |
| Borealis | 33.63 | 74.10 | 24.02 |
| Sceaux | 48.87 | 75.76 | 53.13 |
| SMART | 48.87 | 75.76 | 35.84 |
| IDEAL | 40.92 | 102.31 | 16.71 |

Table 1 shows the different metrics used for forecasting by Chronos model.

Table 2: Number of default buildings with NRMSE error more than 50 from both categories.

| Dataset | #Buildings | #High-error Buildings |
|---|---|---|
| BDG-2 | 611 | 17 |
| Buildings-900K-test | 565 | 115 |
| Electricity | 359 | 14 |
| LCL | 713 | 1241 |
| BOREALIS | 15 | 13 |
| SCEAUX | 4 | 4 |
| SMART | 11 | 11 |
| IDEAL | 350 | 343 |

Table 2 summarizes the buildings from commercial and residential buildings having NRMSE error more than 50.

Table 3 summarizes the buildings from each site of BDG-2 dataset having NRMSE error more than 50. It shows that despite each site being in different location, model able to capture the pattern and forecast accurately.

Table 3: Number of default buildings per site with NRMSE error more than 50 for BDG-2 dataset.

| **Site** | **Location** | **#Buildings** | **#High-error Buildings** |
|---|---|---|---|
| Bear | Berkeley, CA | 169 | 4 |
| Fox | Tempe, AZ | 265 | 5 |
| Rat | Washington DC | 539 | 6 |
| Panther | Orlando, FL | 105 | 2 |