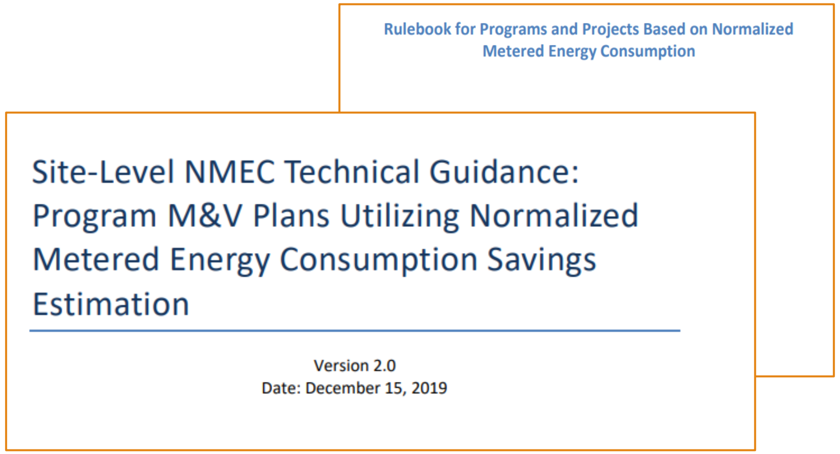
NMEC Program Design with Missing Data, Zero Values, and Differing Meter Resolutions

Summary and Findings

- Allow for building energy model flexibility
  - Model structure (algorithm)
  - Variable inputs
  - Daily aggregation (interval data)
  - Stretching baseline and/or performance period timelines
- Industry accepted model metrics and final savings uncertainty calculations may not always align
- Allow operational time to vet available non-weather model inputs
- Set program participation data quality thresholds

Background

- Program: Central Water Heater Multifamily Building Solution Program
  - Utility sponsor – Southern California Gas (SCG)
  - Initiated in 2016
- Enabling Regulatory framework
  - 2015 California Assembly Bill (AB) 802 enabled program
  - Program advanced as a High Utility Project or Program (HOPPs)
  - Advice letter 4965-A was approved by California Public Utilities Commission (CPUC) August 2016
- Normalized Metered Energy Consumption (NMEC) program design
  - Novel data driven program design leveraging gas AMI
  - Similar in scope to other whole building program designs (SEM, P4P)
    - . . . with weather normalization expectation affecting
  - Program launched prior to now robust published CA NMEC guidelines!
    - Program methodology approved via advice letter in lieu of deferring to published guidelines


Rulebook for Programs and Projects Based on Normalized Metered Energy Consumption

Site-Level NMEC Technical Guidance: Program M&V Plans Utilizing Normalized Metered Energy Consumption Savings Estimation

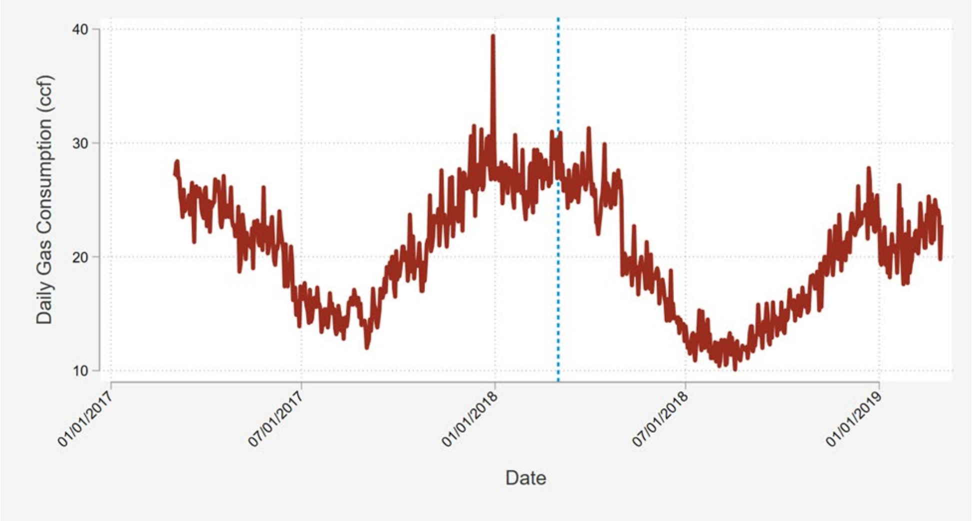
Version 2.0
Date: December 15, 2019

Program Design

- Program eligibility
  - Natural gas heated multifamily buildings
  - Master metered with installed AMI (hourly interval)
  - Built prior to 1984
  - Greater Los Angeles SCG service area
  - Initial 20 program sites
- Program Measures
  - Central storage water heater or boilers
  - Central water heater modulating temperature controllers
  - Hot water system usage monitoring
  - Low flow showerheads and faucet aerators
  - Circulating demand pumps with controllers
- Data requirements
  - 12 months pre (baseline) and post (performance) project interval hourly gas readings
  - Final savings and incentives based solely on whole building prediction modeling

M&V Methodology

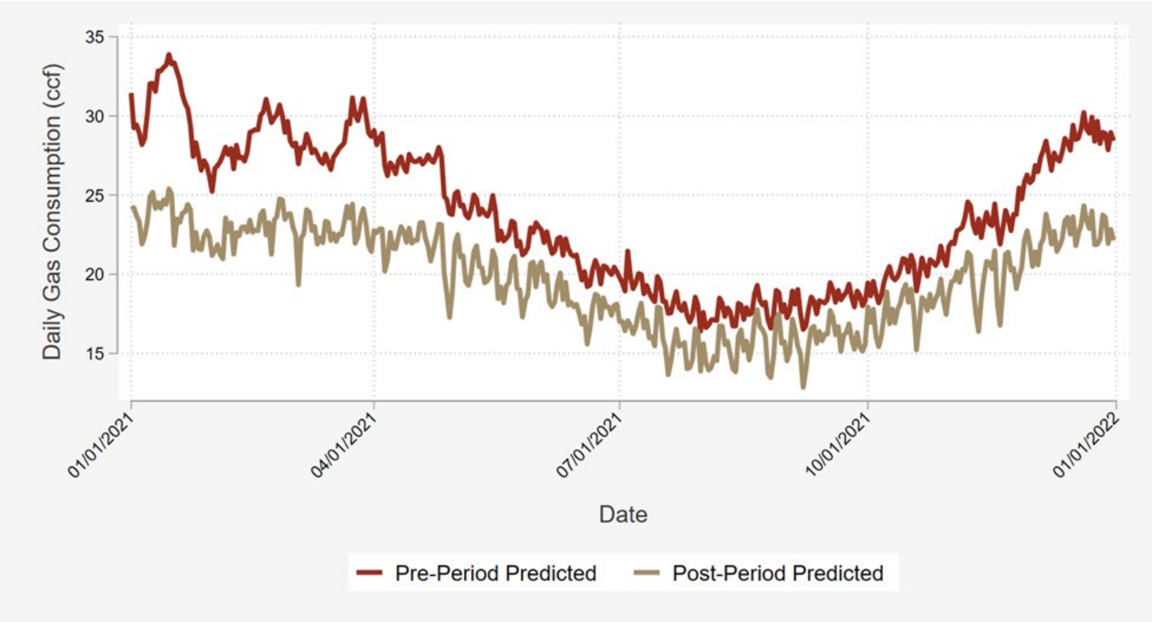
- Baseline/Performance Period Models
  - Hourly precision
  - TOW method characteristics – 2011 LBNL Published DR model method
    - Each hour of the week (n=168) is a separate data feature (variable)
    - Temperature – Seasonal/annual non-linear relationship between energy captured by linear spline features
      - 6 variables representing temperature buckets <20, 20-39, 40-59, 60-79, 80-100, 100+
    - Separate models for occupied/unoccupied time hours of the week*
  - Other included model variables
    - Heating degree hour moving average
    - Holiday indicator
  - Other variable research
    - Water and occupancy
  - Ordinary least squares regression (OLS)



*Not incorporated in final M&V method, because of constant multifamily building occupancy
**Current CA NMEC program guidelines call for CA specific normal climate datasets

M&V Methodology cont. . .

- NMEC – What is weather normalization?
  - Requires baseline and performance period energy prediction models
  - Final savings models projected using Typical Meteorological Year (TMY) datasets*
  - Performance Period Normalized Savings = [baseline forecast] – [performance forecast]
  - Removing short term weather effects from savings facilitates resource plan incorporation
  - Other variables (if known) can be normalized



*CA currently supports state based normalized weather datasets

M&V Methodology cont. . .

- Modeling Metrics (baseline and performance)
  - Variability (accuracy) – CV(RMSE)
    - Average model miss scaled by average hourly gas usage
    - Target threshold - < 25%
  - Bias – Net Bias Error (NBE)
    - Is the model more likely to miss high or low?
    - Target threshold – between -.5% and .5%
  - Explained variance – R^2
    - How well do your prediction variables explain hourly gas usage
    - Target threshold – > 70%

$$CV(RMSE) = \frac{1}{\bar{y}}\left[\frac{\sum(y_i - \hat{y})^2}{(n-p)}\right]^{1/2}$$

$$NBE = 100 * \frac{\sum_i (E_i - \hat{E}_i)}{\sum_i E_i}$$

- Savings Uncertainty
  - Fractional Savings Uncertainty (FSU)
    - Savings confidence interval adjusted for correlation between hourly data points. Divided by total savings estimate
    - Target threshold - < 50%

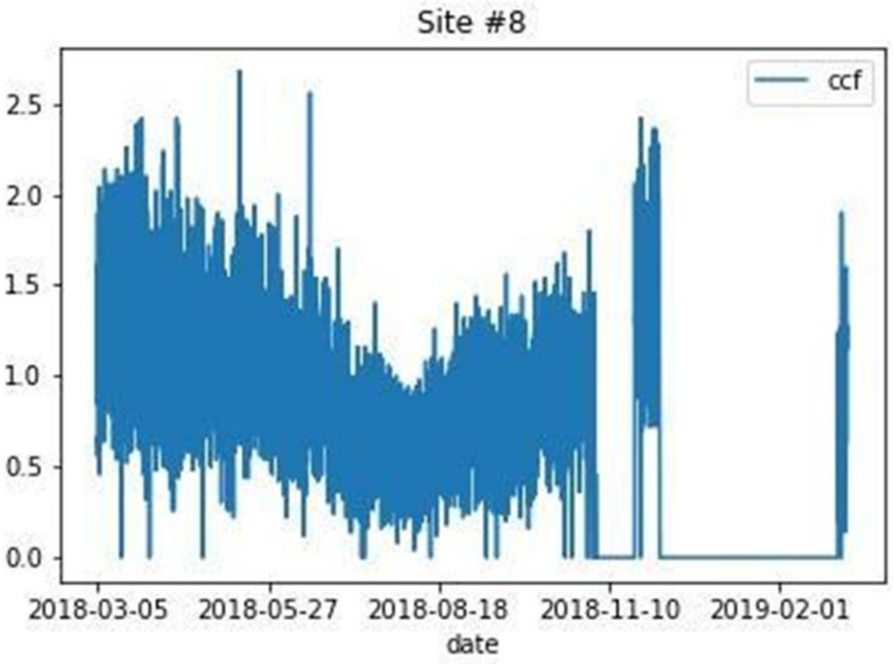
$$Uncertainty = t \times 1.26 \times CV(RMSE) \times \sqrt{\frac{n}{n'} * \left(1 + \frac{2}{n'}\right) * \frac{1}{m}}$$

Data Quality Issues

- Pre-screening gas AMI data quality not a program requirement
- Utility reviewed all data quality concerned sites and verified data accuracy
- Four identified data issues were identified during data pre-screening
  - Zero-value (CCF) reads
    - May miss key usage seasons
  - Poor temperature and gas usage correlation
    - Mild climate – Greater Los Angeles area
    - Isolated domestic hot water usage measured end use
  - Low gas usage variability (same value repeated)
    - Impacts modeling capabilities
  - Low meter resolution
    - Hourly data exported as integers – meter programming

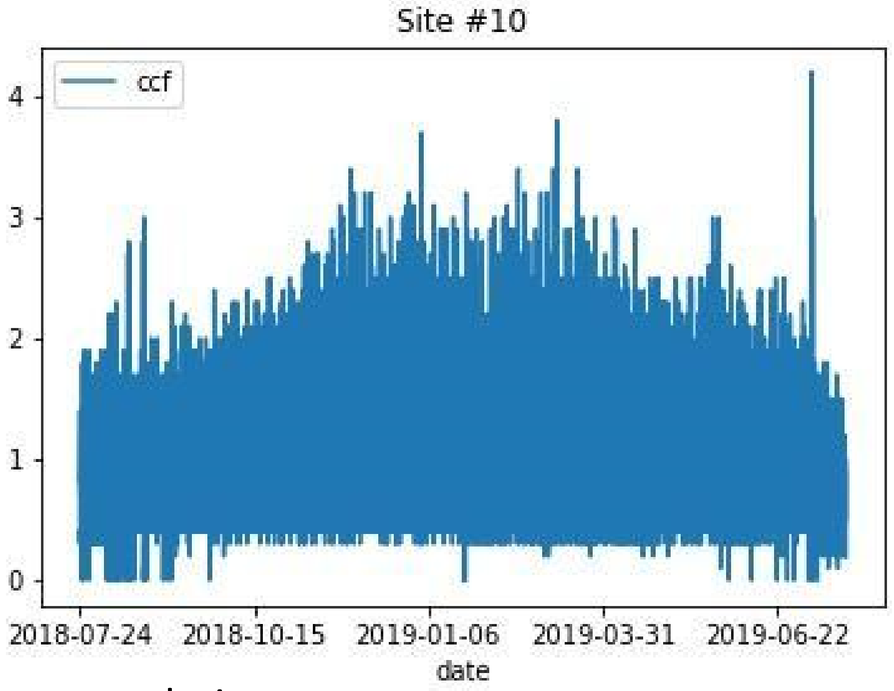
Data Quality Issues – Zero reads

- Resolution – remove day when 22+ consecutive missing
- In extreme case go back and pull more data



Data Quality Issues – Low temp-gas correlation

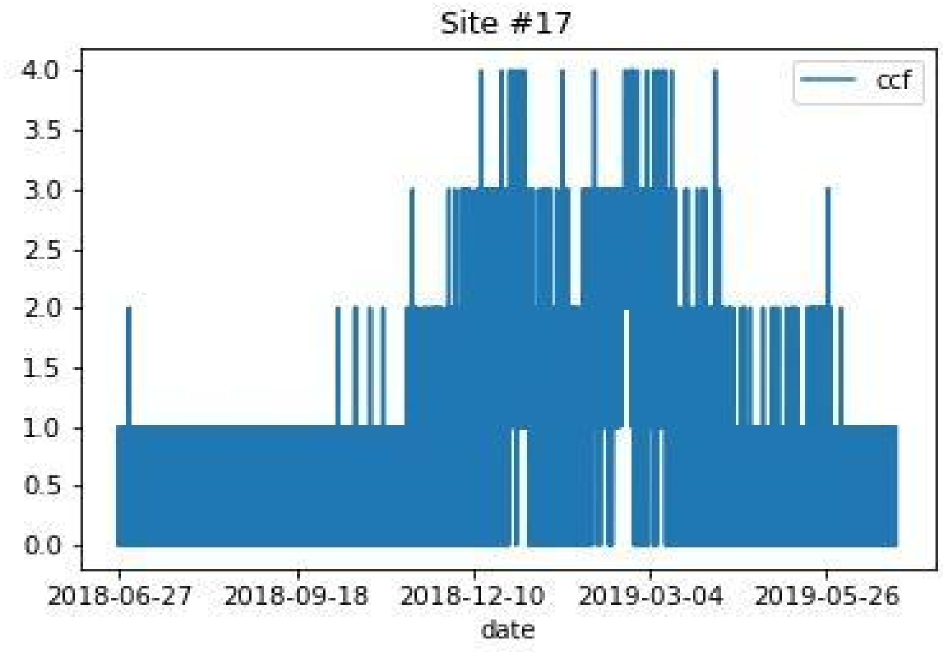
- Resolution – consider daily model instead of hourly*, find additional model variables
- Hourly variability masks seasonal correlation in model building



*Post-mortem recommendation

Data Quality Issues – Low gas variability/meter resolution

- Resolution – consider daily model instead of hourly*
- Low gas usage variability makes model building challenging



*Post-mortem recommendation

Modeling Metric Results

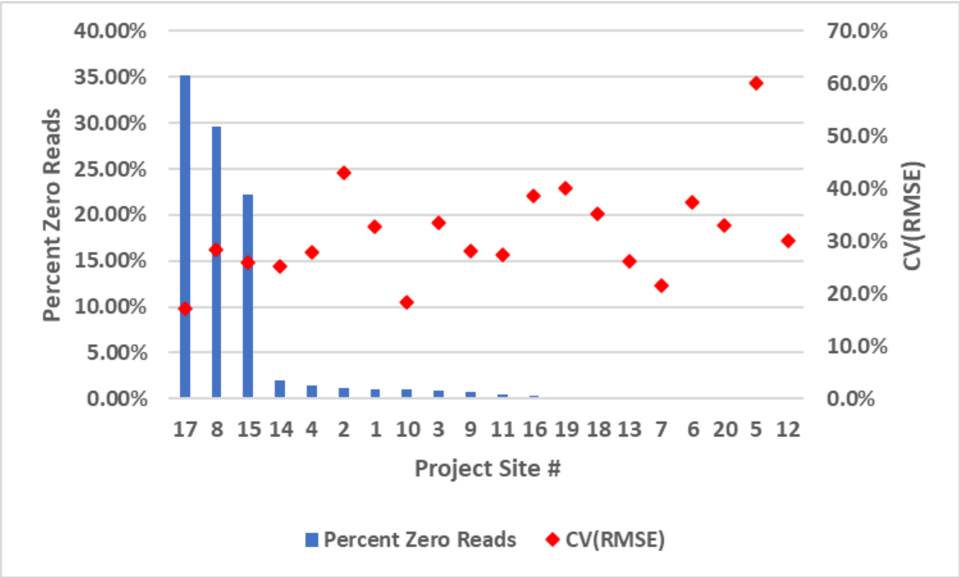
- 80% of program sites failed CV(RMSE) and R^2 program thresholds for baseline and performance models

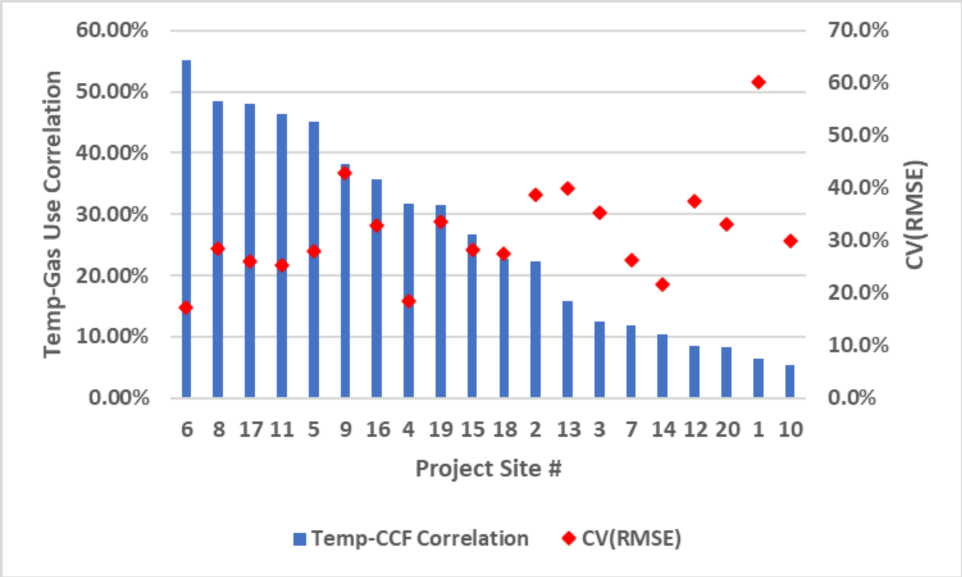
while. . .

- 20% of program sites passed model uncertainty thresholds
- Failed model goodness of fits were not good predictors of savings uncertainty

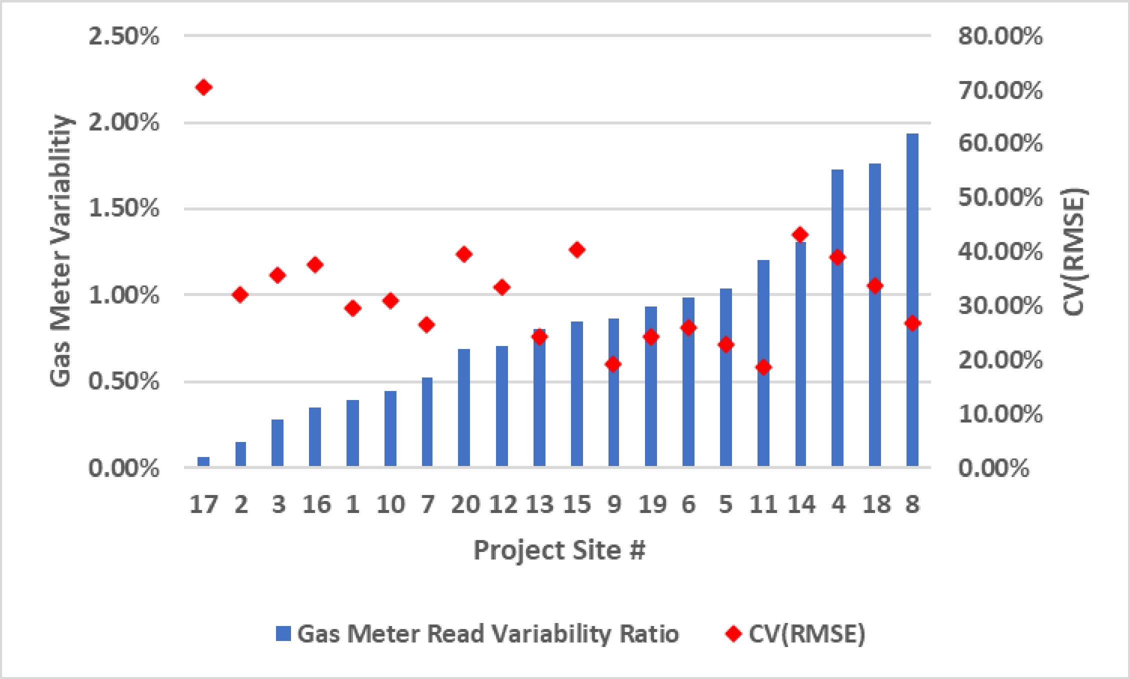
	# Sites Failing R^2 (< 70%)	# Sites Failing CV(RMSE) (>25%)	# Sites Failing FSU (< 0 or > 50%)	Avg Site Savings (% baseline usage)
Baseline	16	15	4	8.7%
Performance	16	14	4	8.7%

Data quality issues and model metrics





Data quality issues and model metrics



Alternative Modeling Approaches

- Do alternative model algorithms improve metrics? No, not much
  - Tested 9 advanced regression and various machine learning model types on 4 poorest model sites
  - Only slight metric improvements – Poor data fit is simply a poor data fit

- Tested daily aggregation in combination with alternative model option. Yes.
  - Hourly to daily gas usage aggregation help correct for data quality issues
  - 4 poorest model fit site were all able to pass CV(RMSE) metrics using daily models

Program Site	Hourly Model CV(RMSE)	Model Type	Daily Energy CV(RMSE)
17	70%	Gradient Boosting	24%
14	43%	Ridge regression	15%
15	40%	Ridge regression	23%
20	40%	Random forest	11%

Recommendations

- Program rules, designs, and planning can help hedge potential data quality issues
  - Allow time for site/program data exploration for non-weather covariates (water, occupancy, etc.)
  - Build in data quality screens into program requirements
  - Allow model type flexibility
    - Advanced model types will not rescue
  - Be willing to give up hourly granularity (e.g. daily aggregation) to improve model metrics and overcome data quality issues
- Industry standard (ASHRAE, IPMVP) baseline model metrics are not always good predictors of model uncertainty thresholds
  - Consideration for future IPMVP standard updates – future research
  - Consider data science best practices instead of relying solely on traditional statistical metrics
    - E.g. cross validation, train-test methods, re-sampling model metric calculations