Weather and Irrigation Normalization Technical Conference

June 19, 2008

Agenda

- Study of the methodology proposed by PNGC
  - Review concerns with models
  - Review of results to date
  - Continuation of efforts?
- Next Steps?

PNGC Model –Generalization

- Models loads-less large industrial
- Models with several weather variables
  - (CDD,CDD lags HDD lags, Prec accumulations)
- Models to account for Commercial, Residential, and Irrigation changes
  - Several Economic and Demographic factors
- Models all PNGC customers collectively

Concerns with PNGC Model

- Great model for forecasting not necessarily for weather normalizing
- Models all customers collectively
  - Not a consistent approach for all customers in Agency
  - Could average effects of customers into single impact
- Complicates the process unnecessarily
  - Extra variables may be good for forecasting but are not necessary for weather normalization
  - May be cross-correlations in variables that impact coefficient estimates
  - Only correctly estimated weather coefficients are needed for normalizing
- Simplifies some weather variables
  - No review of CDD and HDD on a degree by degree basis

Review step 1

- Duplicate PNGC model at Agency
  - 60255 observations – 15 customers, 11 year of daily date
  - 50 independent variables
  - R-squared .9733
  - Root MSE 136375

Review step 2

- Separate into model for each customer
  - Selected customer 8
    - 4017 observations –11 years of daily date for 1 customer
    - 50 independent variables
    - R-squared .9957
    - Root MSE 43985
    - **(Model failed) Model is not full rank. Least-squares solutions for the parameters are not unique. Some statistics for the model will be misleading.**

Review step 3

- Simplify separate model for each customer
  - Selected customer 8
    - 4017 observations –11 years of daily date for one customer
    - 28 independent variables (mostly indicator variables)
      - 9 yearly variables, 11 monthly variables, 6 day type variables
      - 2 weather variables Cdd_60 Hdd_55
    - R-squared .9503
    - Root MSE 43444
    - Cdd coefficient 2094.05
      - Standard Error 281.48
      - T-stat 7.44
    - HDD coefficient 19371
      - Standard Error 173.04
      - T-stat 111.94

Review step 4

- Refine weather variable
  - If weather is not correctly modeled then coefficient will be wrong.
    - If correct relationships is KWH= 2 * hdd_base 68 and modeled with hdd_base 65 then coefficient will be too high or error to great.

  Process is to model in 1 degree increments for HDD and CDD and stop when MSE of model is lowest. Should also give best T-stat and R-squared

Review step 4

- Refine CDD variable
  - Selected customer 8
    - 4017 observations –11 years of daily date for one customer
    - 28 independent variables (mostly indicator variables)
      - 9 yearly variables, 11 monthly variables, 6 day type variables
      - 2 weather variables Cdd_68 Hdd_55
    - R-squared .9521
    - Root MSE 42784
    - Cdd coefficient 7500.56
      - Standard Error 557.34
      - T-stat 13.46
    - HDD coefficient 19327
      - Standard Error 170.37
      - T-stat 113.44

Review step 5

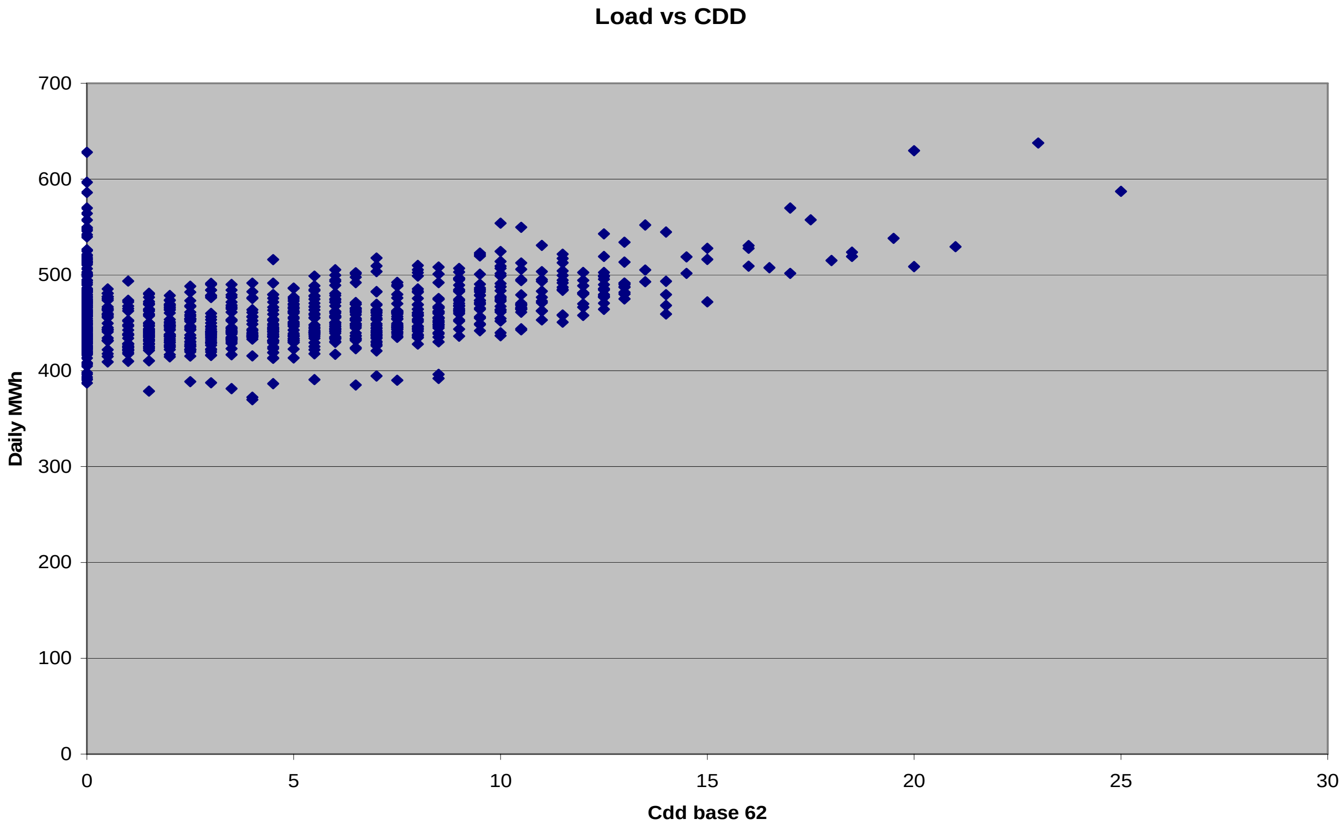
- Refine HDD
  - base 54
    - 4017 observations –11 years of daily date for one customer
    - 28 independent variables (mostly indicator variables)
      - 9 yearly variables, 11 monthly variables, 6 day type variables
      - 2 weather variables Cdd_68 Hdd_54
    - R-squared .9320
    - Root MSE 48207
    - Cdd coefficient **0.**
      - Standard Error 557.34
      - T-stat 13.46
    - HDD coefficient 19408
      - Standard Error 199.76
      - T-stat 97.15
    - **(Model failed) Model is not full rank. Least-squares solutions for the parameters are not unique. Some statistics for the model will be misleading.**

Review step 5

- Find out what went wrong-why did CDD coefficient go to zero?
  - Separate out summer months to model separately
    - May- September (no better results)
    - June-September (no better results)
    - July-September (no better results)
  - Why?

Review step 5

- July-September data reveals why
  - Modeling fitting a straight horizontal line



Review step 6

- Refined CDD- using summer only data
- Summer defined as July 21-Sept 10 of each year
  - base 68 degrees
    - 583 observations –11 years of daily date for one customer summer season
    - 13 independent variables (mostly indicator variables)
      - 9 yearly variables, 3 day type variables
      - 1 weather variables Cdd_68
    - R-squared .7130
    - Root MSE 16184
    - Cdd coefficient 7228.34
      - Standard Error 246.26
      - T-stat 29.35
- Note sizeable change in the R-square and coefficient

Review step 7

- Refined CDD- using summer only data
- Summer defined as July 21-Sept 10 of each year
- Step through the CDD in 1 degree increments to see change in results
- Several model check to select optimum

Cdd base	Coefficient	Root MSE	Adjusted R-squared	St Error	T-Stat	R-squared
64	4963.06	16622	0.6904	176.88	28.06	0.6973
66	5918.86	16039	0.7117	198.66	29.79	0.7182
67	6512.12	16006	0.7129	217.83	29.89	0.7193
68	7228.34	16184	0.7065	246.26	29.35	0.713

Review step 7

- Refined CDD- using summer only data
- Summer defined as July 21-Sept 10 of each year
  - base 67 degrees (best fit)
    - 583 observations –11 years of daily date for one customer summer season
    - 13 independent variables (mostly indicator variables)
      - 9 yearly variables, 3 day type variables
      - 1 weather variables Cdd_67
    - R-squared .7193
    - Root MSE 16006
    - Cdd coefficient 6512.12
      - Standard Error 217.83
      - T-stat 29.89

Summary to date

- PNGC model appears to give cross correlations and less precise weather coefficients.
  - Model works well for forecasting
  - Model coefficients are not precise as necessary for weather normalization
- As indicated in prior technical workshops
  - Splitting the data into seasons will be required.
  - Individual customer models will be required.