Marquette Energy Analytics

Improving Natural Gas Demand Forecasting Through the Reconciliation of Incoherent Data Hierarchies

July 3, 2024

Colin Quinn
Richard Povinelli

info@marquetteenergyanalytics.com
(414) 765-2839

309 N. Water St.
Suite 400
Milwaukee, WI 53202

International Symposium On Forecasting 2024

Colin Quinn

Milwaukee, Wisconsin
Marquette University
Computer Science
Applied Statistics

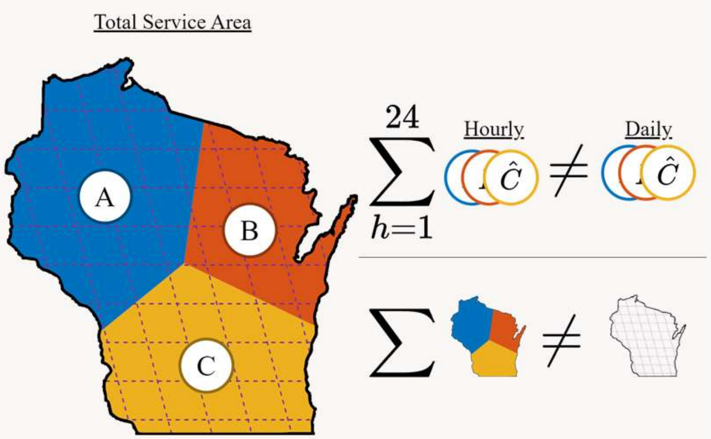
Marquette Energy Analytics

Agenda

- Natural gas demand forecasting and reconciliation
  - Coherent forecasts, aligned decision making
  - Natural gas consumption data organized hierarchically
- Challenges in natural gas consumption data
  - Billing cycles, data frequency
- Hierarchical preprocessing technique for data incoherence
  - Implications for the natural gas industry
- Case study results and analysis
  - Cross-sectional reconciliation of natural gas demand forecasts
- Summary and future directions

Natural gas demand forecasting and reconciliation

- Natural gas is a fossil fuel energy source extracted for sale and consumption
  - Residential, commercial, and industrial uses
- Gas consumption data can be hierarchically organized to improve demand forecasting



Coherent gas demand forecasts

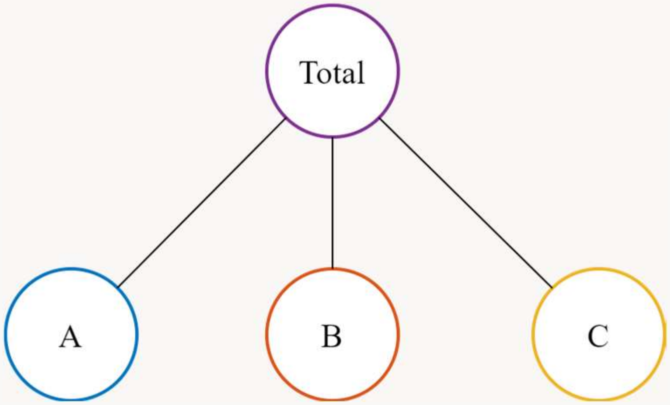
via cross-sectional reconciliation

Investigate tradeoff:

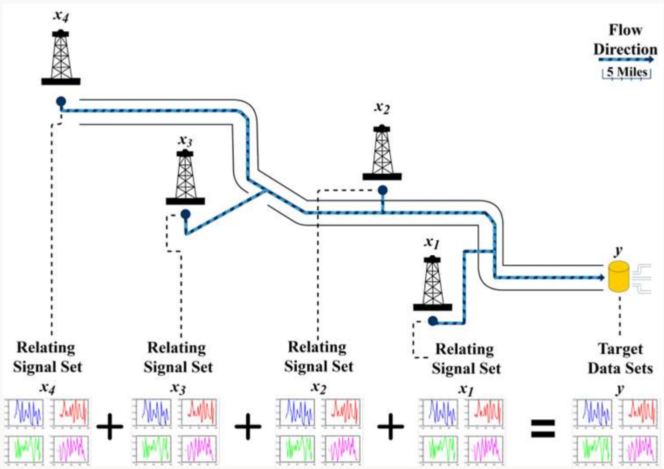
Coherence VS "Accuracy"

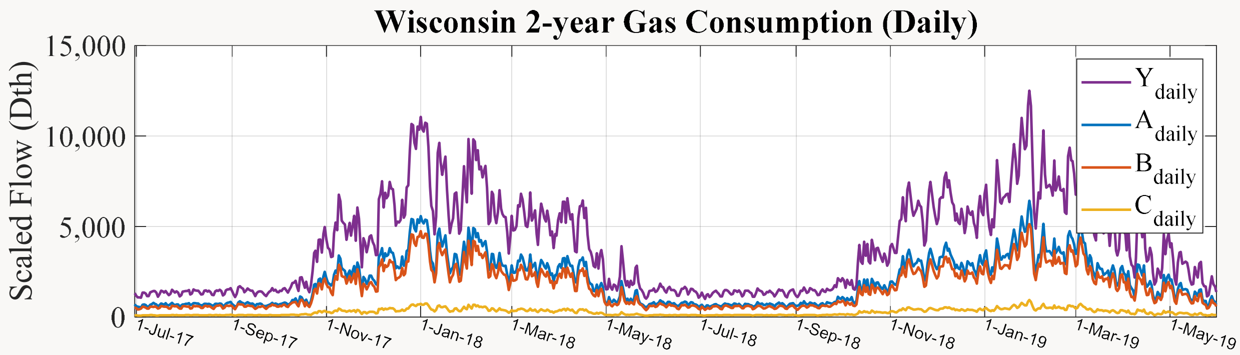
Natural gas consumption data organized hierarchically

- Three approaches to organizing data hierarchically
  1. Cross-sectional
  2. Temporal
  3. Cross-temporal
- Gas consumption data organized geographically, temporally
  - Locked to the time of analysis



State of Wisconsin divided into service areas

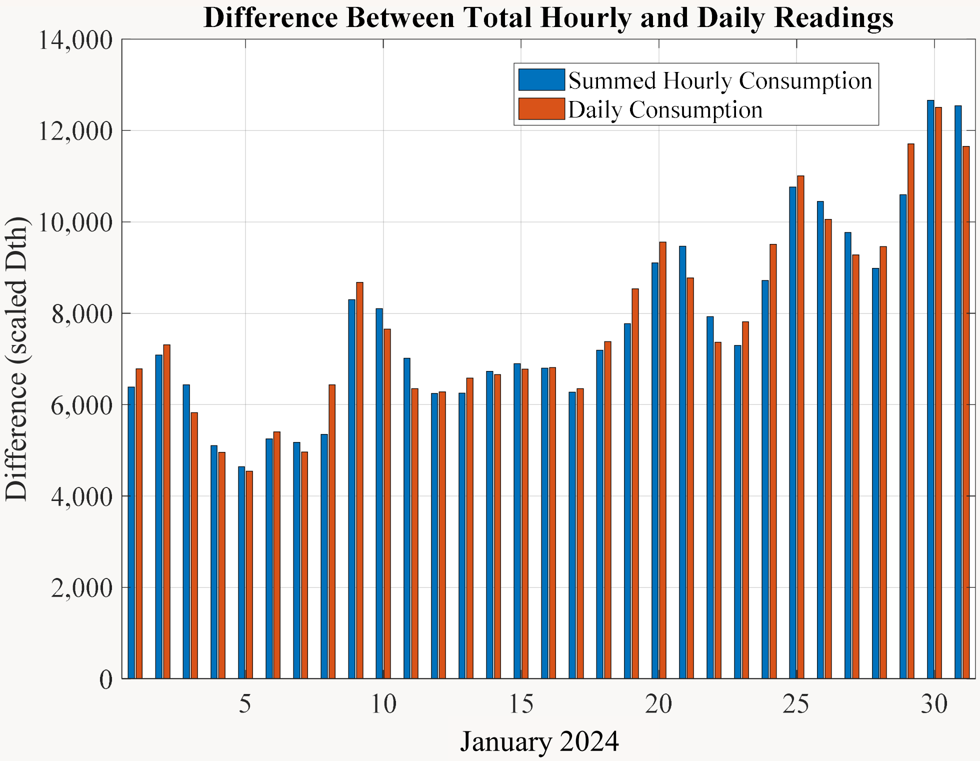


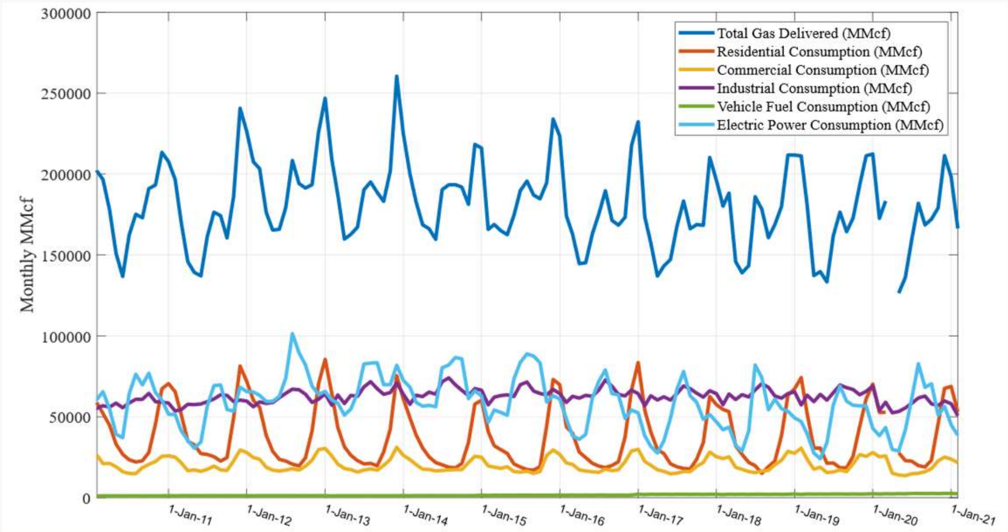


- Challenge: forecasts at individual levels rarely align and provide conflicting intra-level results

Natural gas consumption data challenges

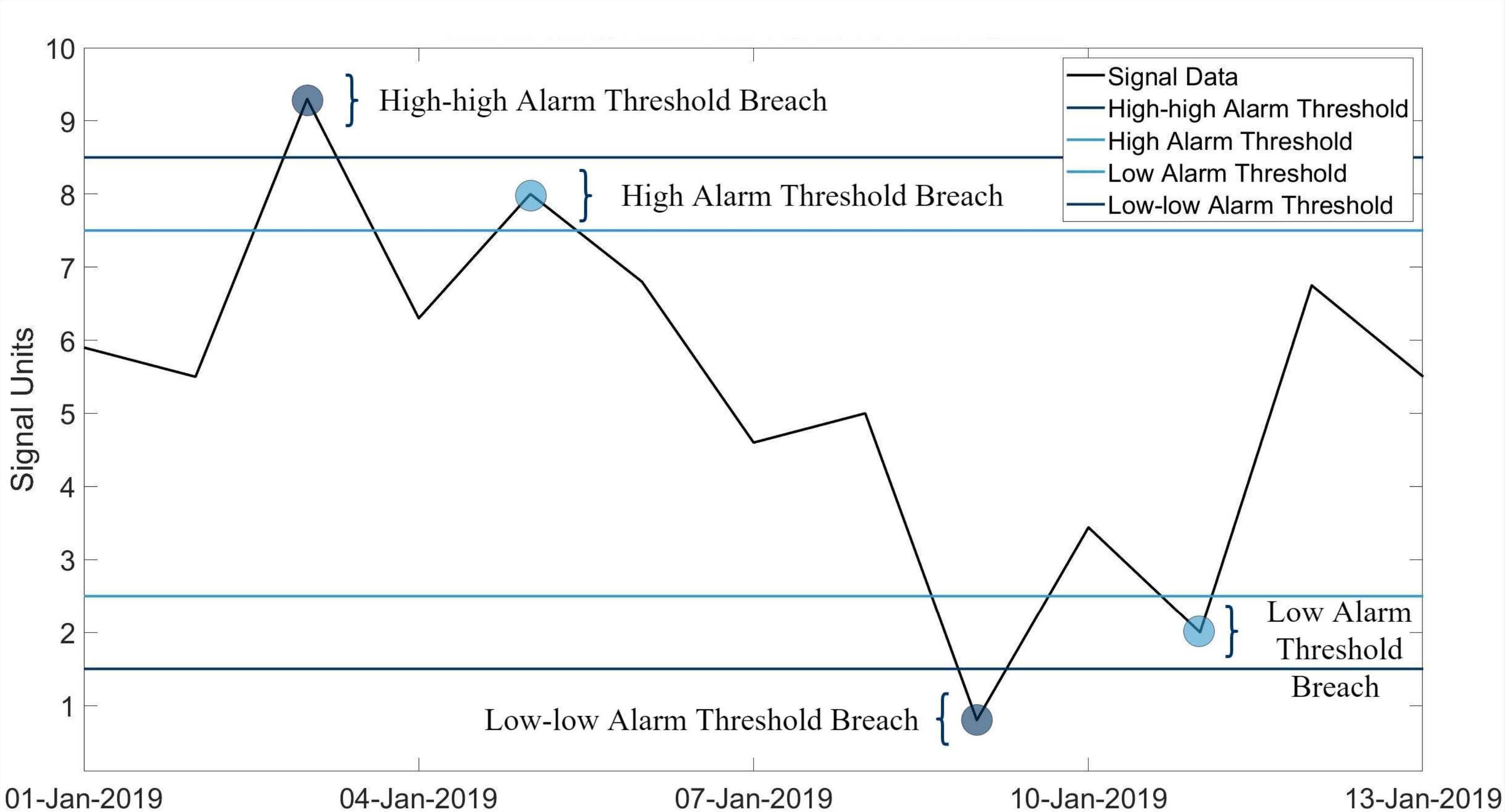
- In general, observed gas consumption data are not **coherent**



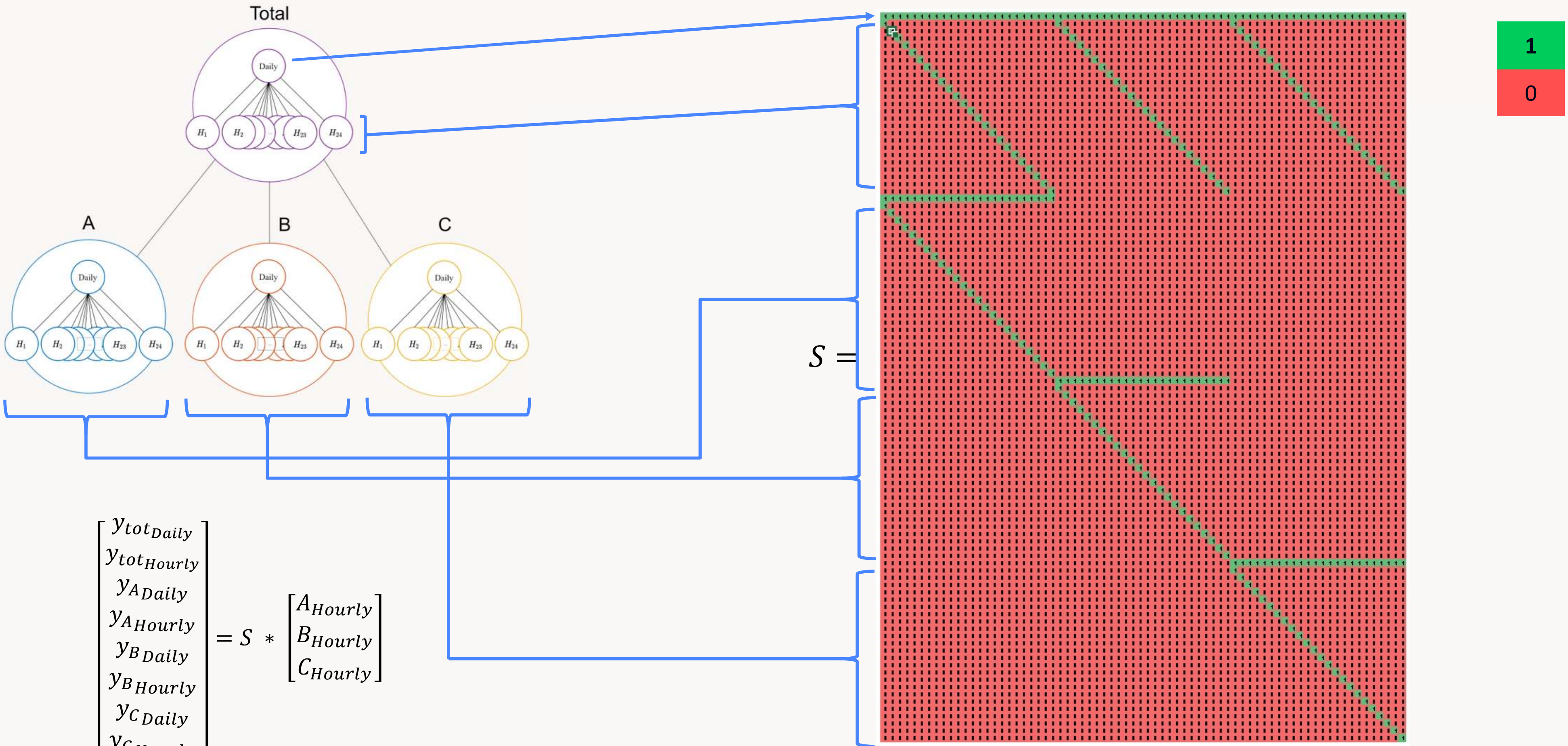


- **Forecast reconciliation** is the process of adjusting forecasts to make them coherent

Example: Alarm forecasting

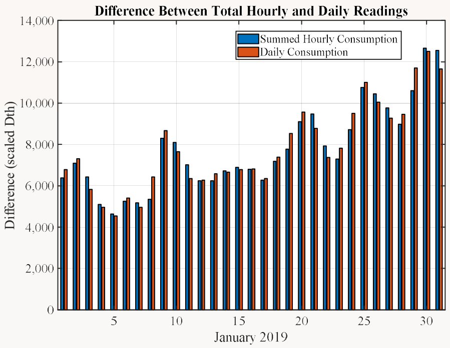


Cross-temporal hierarchy for gas demand forecasting



$$\begin{bmatrix} y_{tot_{Daily}} \\ y_{tot_{Hourly}} \\ y_{A_{Daily}} \\ y_{A_{Hourly}} \\ y_{B_{Daily}} \\ y_{B_{Hourly}} \\ y_{C_{Daily}} \\ y_{C_{Hourly}} \end{bmatrix} = S \ast \begin{bmatrix} A_{Hourly} \\ B_{Hourly} \\ C_{Hourly} \end{bmatrix}$$

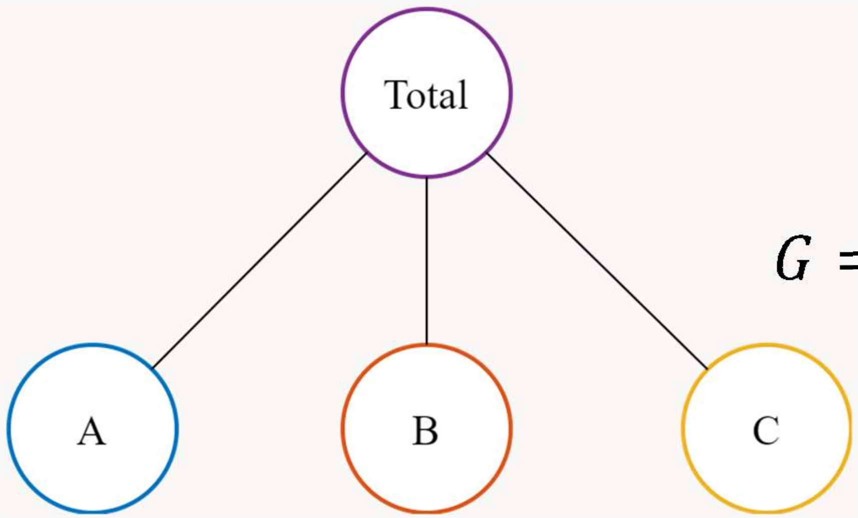
Preprocessing technique for data incoherence



- No ground truth
  - Relevant information available for demand estimation
  - Customer → product owner
- Align incoherent base forecasts according to predefined linear constraints
  - Original hierarchical time series reconciliation → identified incoherence by comparing aggregated versions of time series
  - Lost and unaccounted for gas (LAUF)
- This study aims to investigate the effectiveness of a weighted reconciliation preprocessing technique applied to natural gas consumption data with significant in-sample incoherence to improve out-of-sample base forecast accuracy

Preprocessing technique for data incoherence

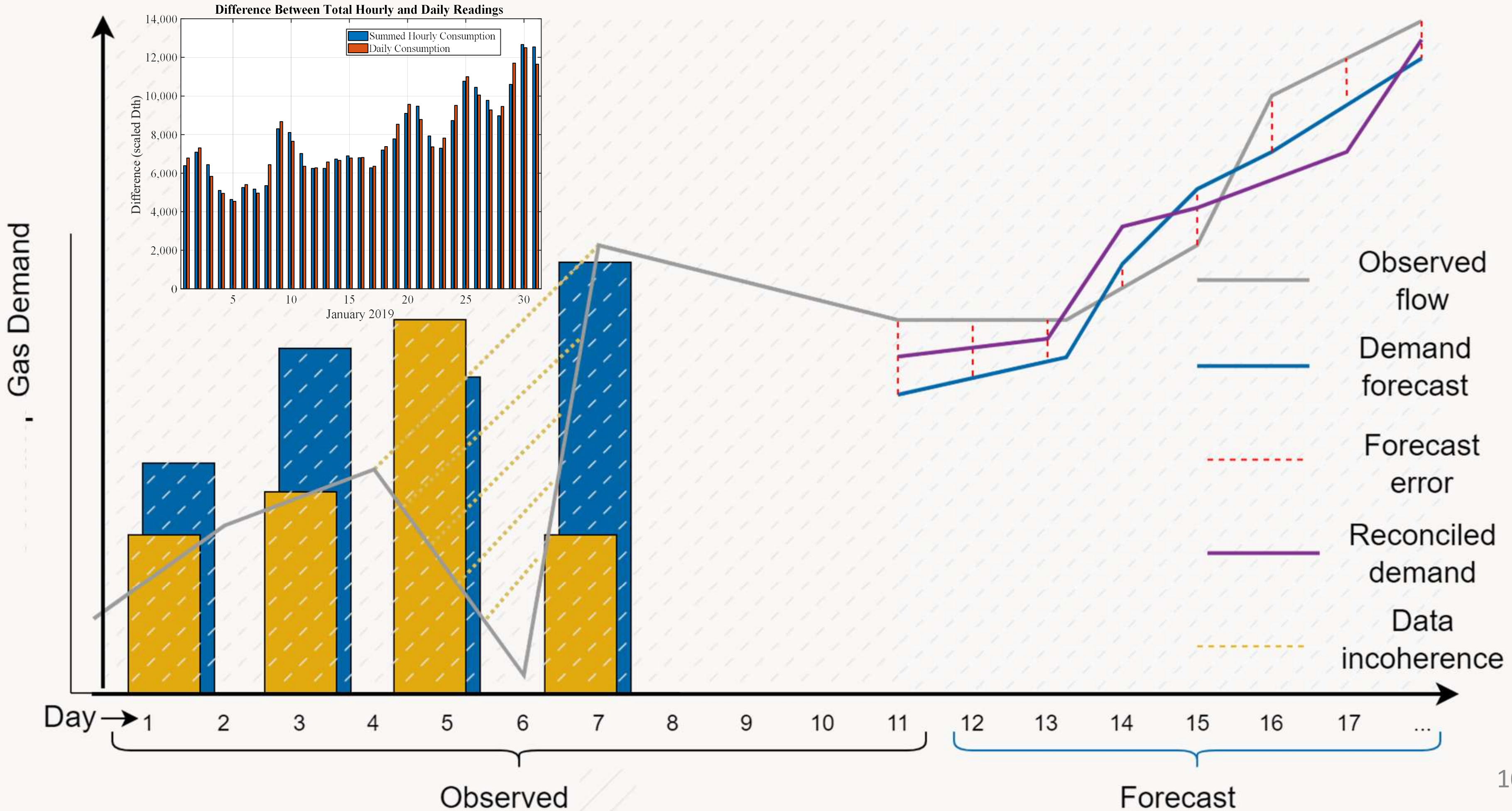
- Optimal forecast reconciliation (Hyndman et al., 2019)
  - Minimizes the forecast error of the set of coherent forecasts
  - $b = (A, B, C)'$
  - $y = (y_{tot}, y_A, y_B, y_c)$
  - $y = Sb$
- Summation matrix
- $S = \begin{matrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$
- $\hat{y}_h$: h step ahead forecasts for y



$$\tilde{y}_h = SG\hat{y}_h$$

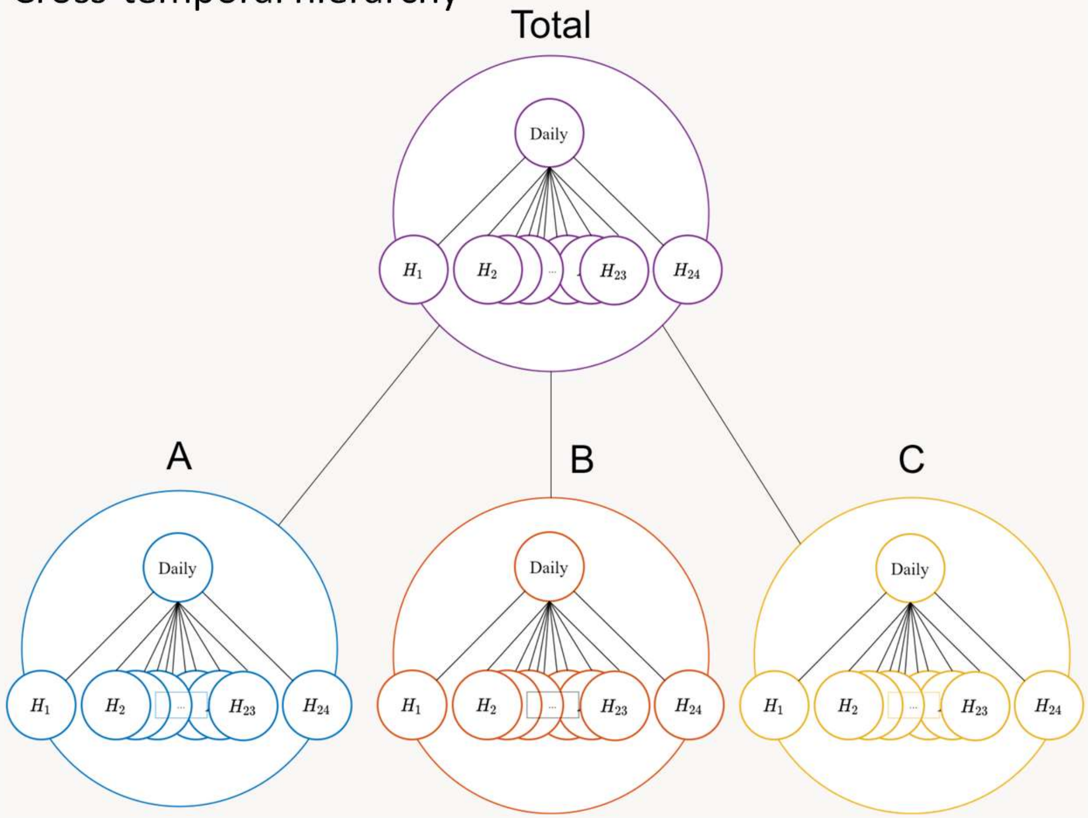
$$G = \left(S'W_h^{-1}S\right)^{-1}S'W_h^{-1}$$

Preprocessing technique for data incoherence



Hierarchy for gas demand forecasting

- Forecast coherency in a natural gas distribution company setting
  - Cross-temporal hierarchy



$$b = (A_{Hourly}, B_{Hourly}, C_{Hourly})'$$

$$y = Sb$$

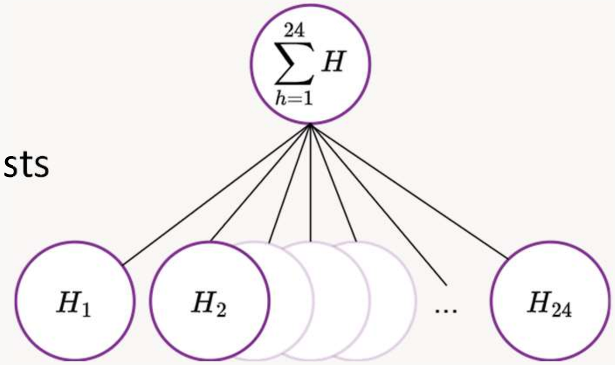
$$\begin{bmatrix} y_{tot_{Daily}} \\ y_{tot_{Hourly}} \\ y_{A_{Daily}} \\ y_{A_{Hourly}} \\ y_{B_{Daily}} \\ y_{B_{Hourly}} \\ y_{C_{Daily}} \\ y_{C_{Hourly}} \end{bmatrix} = S * \begin{bmatrix} A_{Hourly} \\ B_{Hourly} \\ C_{Hourly} \end{bmatrix}$$

Preprocessing technique for data incoherence

$$\tilde{y}_t = SG\hat{y}_t$$

$$G = (S'W_t^{-1}S)^{-1}S'W_t^{-1}$$

- Tradeoff: coherence and accuracy
  - G – ideally small permutation
- Advantages
  1. Point forecasts are reconciled across all levels of the hierarchy
     - In our case, reconciled across planning horizons and gas service areas
  2. Intra-hierarchy level interactions and correlations taken into account
     - Plan from long- to short-term horizons
  3. Ad hoc adjustments
     - Ensemble base forecasts result with preprocessed reconciled forecasts
  4. Forecast uncertainty



- Construction of S and G is nontrivial and computationally demanding

Case study results and analysis

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{d=1}^{N_D}(\hat{Y}_d - y_d)^2}{N_D}}$$

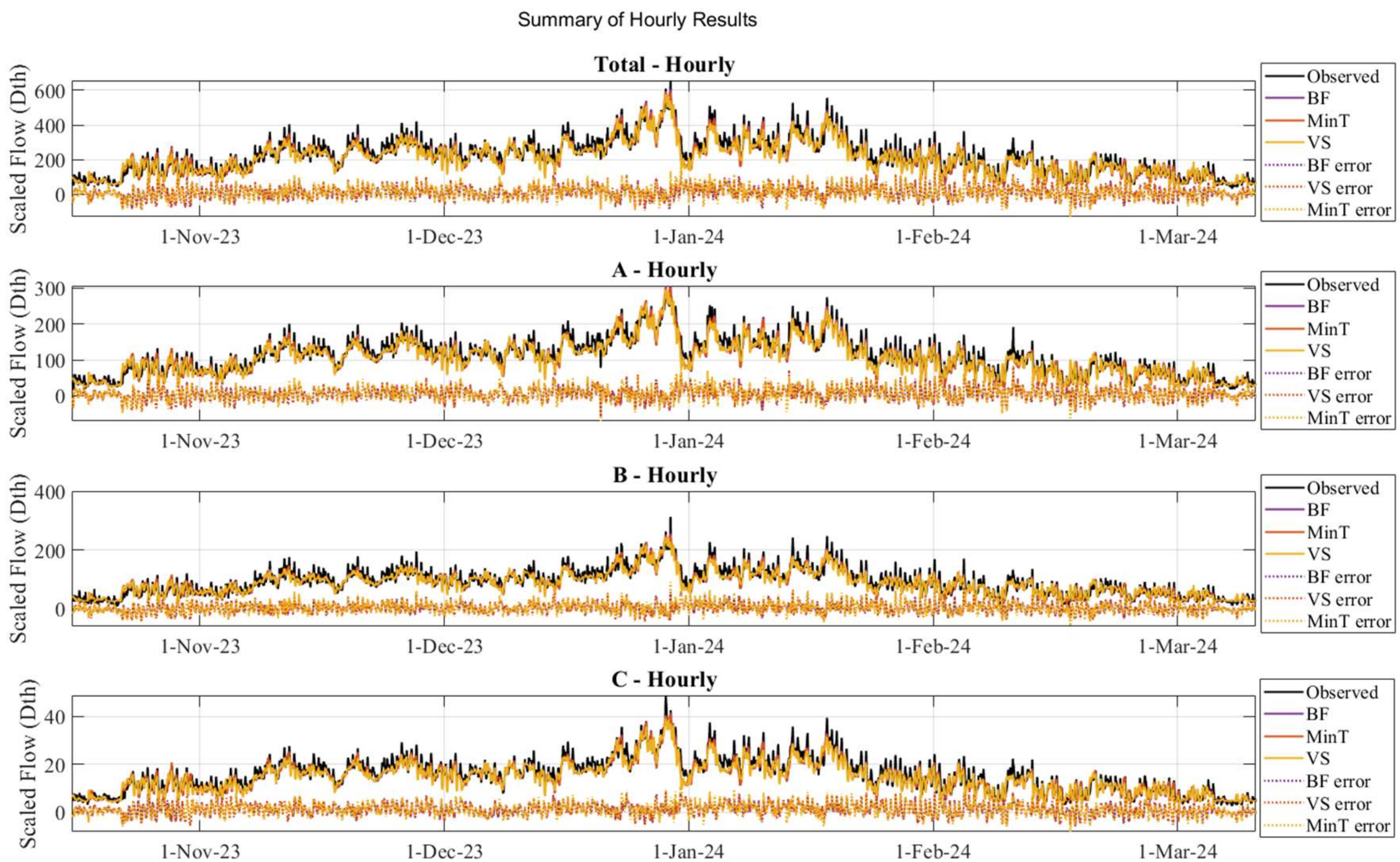
- Forecast accuracy is currently measured as:
  - root mean square error (RMSE)
  - mean absolute percentage error (MAPE)
  - weighed mean absolute percentage error (wMAPE)
- Base model: 5-param numerical weather forecasts

$$\mathrm{MAPE} = \frac{1}{N_D}\sum_{i=1}^{N_D}\frac{|(\hat{Y}_d - y_d)|}{Y_d}.$$

$$\mathrm{wMAPE} = 100\ x\ \frac{\sum_{d=1}^{N^D}|y_d - \hat{Y}_d|}{\sum_{d=1}^{N^D}|y_d|}\%$$

Table 1. Results for Total (daily)

Method	Coherent	RMSE [Dth]	MAPE [%]	wMAPE [%]
NAV	Yes	927.9	16.7	13.8
BASE	No	319.8	7.9	6.2
MinT	Yes	528.3	11.4	9.7
PPHR	No	643.5	10.1	8.4


Summary of Hourly Results
Total - Hourly
Scaled Flow (Dth)
600
400
200
0
1-Nov-23
1-Dec-23
1-Jan-24
1-Feb-24
1-Mar-24
Observed
BF
MinT
VS
BF error
VS error
MinT error
A - Hourly
Scaled Flow (Dth)
300
200
100
0
1-Nov-23
1-Dec-23
1-Jan-24
1-Feb-24
1-Mar-24
Observed
BF
MinT
VS
BF error
VS error
MinT error
B - Hourly
Scaled Flow (Dth)
400
200
0
1-Nov-23
1-Dec-23
1-Jan-24
1-Feb-24
1-Mar-24
Observed
BF
MinT
VS
BF error
VS error
MinT error
C - Hourly
Scaled Flow (Dth)
40
20
0
1-Nov-23
1-Dec-23
1-Jan-24
1-Feb-24
1-Mar-24
Observed
BF
MinT
VS
BF error
VS error
MinT error

Case study results and analysis

	$AvgRelRMSE$	
Unit of Analysis	**Preprocessed**	$\boldsymbol{AvgRelRMSE}$
Temporal	Base	1.000
	MinT	0.973
	Bottom Up	0.986
	Top Down	0.991
Spatial	Base	1.000
	MinT	0.961
	Bottom Up	0.977
	Top Down	0.984
Cross-temporal	Base	1.000
	MinT	**0.927**
	Bottom Up	0.939
	Top Down	0.942

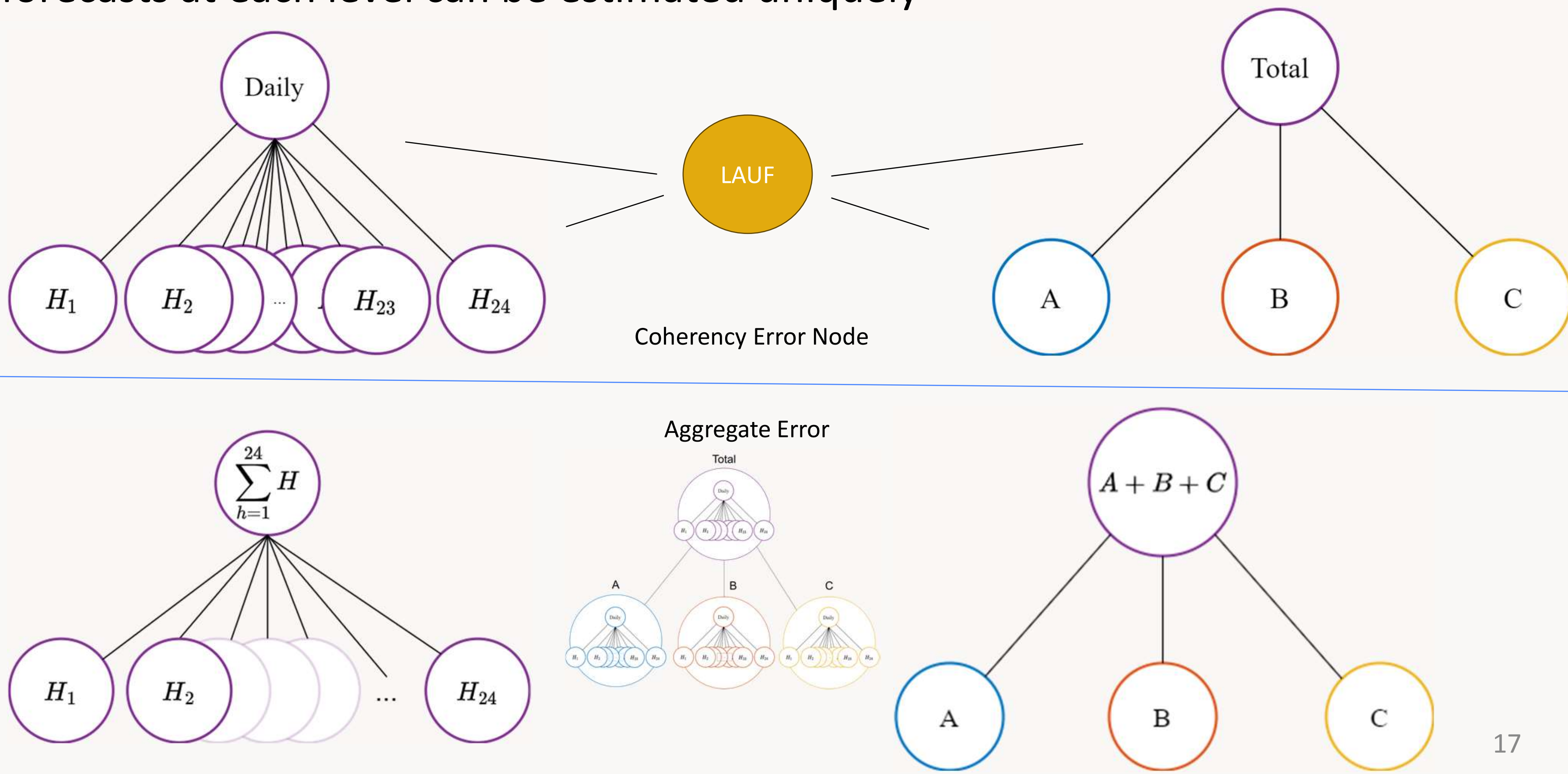
- Total level
- Grouped by Unit of Analysis
- $AvgRelRMSE < 1$
- Cross-temporal approach shows promise
  - Temporal results – averaged geometric mean
- Bottom up and Top Down

Case study results and analysis

	A		B		C		Total		Average	
Resolution	Hourly	Daily	Hourly	Daily	Hourly	Daily	Hourly	Daily	Hourly	Daily
Method					MASE					
BF	3.14	0.69	2.34	0.68	3.11	0.75	2.95	0.68	2.92	0.70
VS	3.16	0.74	2.35	0.75	3.12	0.83	2.92	0.72	2.89	0.76
MinT	**3.10**	**0.66**	**2.29**	**0.65**	**3.06**	**0.73**	2.84	0.63	2.82	0.67
					RMSSE					
BF	2.73	0.60	2.04	0.60	2.81	0.66	2.63	0.60	2.58	0.61
VS	2.75	0.68	2.06	0.68	2.82	0.75	2.61	0.67	2.56	0.69
MinT	2.70	0.58	2.04	0.58	2.79	0.64	2.60	0.57	2.55	0.59
					AMSE					
BF	0.99	0.42	0.26	0.17	0.71	0.31	0.71	0.33	0.67	0.31
VS	1.01	0.42	0.28	0.16	0.72	0.30	0.68	0.30	0.67	0.29
MinT	1.09	0.45	0.33	0.18	0.86	0.36	0.75	0.33	0.76	0.33

Preprocessing technique for data incoherence

- Base forecasts at each level can be estimated uniquely

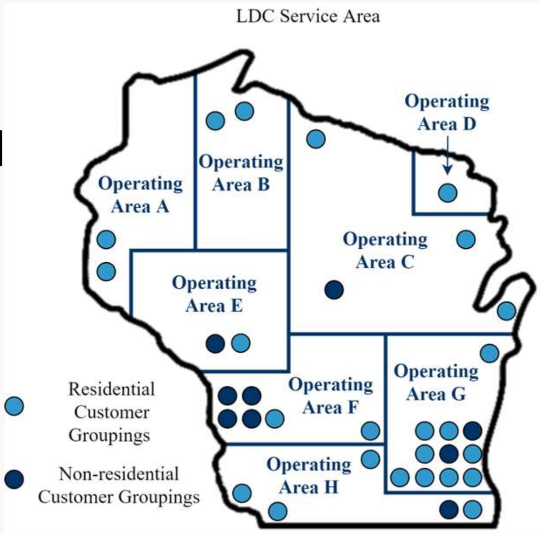


Next steps

- Feasibility of hierarchical forecasting in deep-temporal natural gas demand setting
  - Include more hierarchies (res, non res, …)
  - Estimation of variance-covariance matrix W

- Different levels of gas distribution organization and time horizon of concern
  - Hourly, daily, monthly

- Improve base forecasts over 5-parameter model



Operating Area Code	Number of Customers	Number of Bills
A	82,438	4,946,280
B	25,823	1,549,380
C	147,325	8,839,500
D	50,683	3,040,980
E	13,659	819,540
F	234,752	14,085,120
G	316,166	18,969,960
H	129,154	7,749,240

Questions?

Marquette Energy Analytics

Thank you.

colin.o.quinn@marquette.edu
richard.povinelli@marquette.edu

info@marquetteenergyanalytics.com
(414) 765-2839

309 N. Water St.
Suite 400
Milwaukee, WI 53202