Multi-source Iterative Load Shifting Disaggregation

June 29, 2022

Colin Quinn
Richard Povinelli

info@marquetteenergyanalytics.com
(414) 765-2839

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

International Symposium On Forecasting 2022

Colin Quinn

Milwaukee, Wisconsin

Marquette University

Computer Science

Applied Statistics

Marquette Energy Analytics

Agenda

- Introduction to natural gas consumption, demand, and forecasting
  - Problem Intro
  - H F Overview
  - Method section
  - Results
- Generalized temporal disaggregation
- Multi-source Iterative Load Shifting Disaggregation Algorithm
- Application: case study

Natural gas consumption, demand, and forecasting

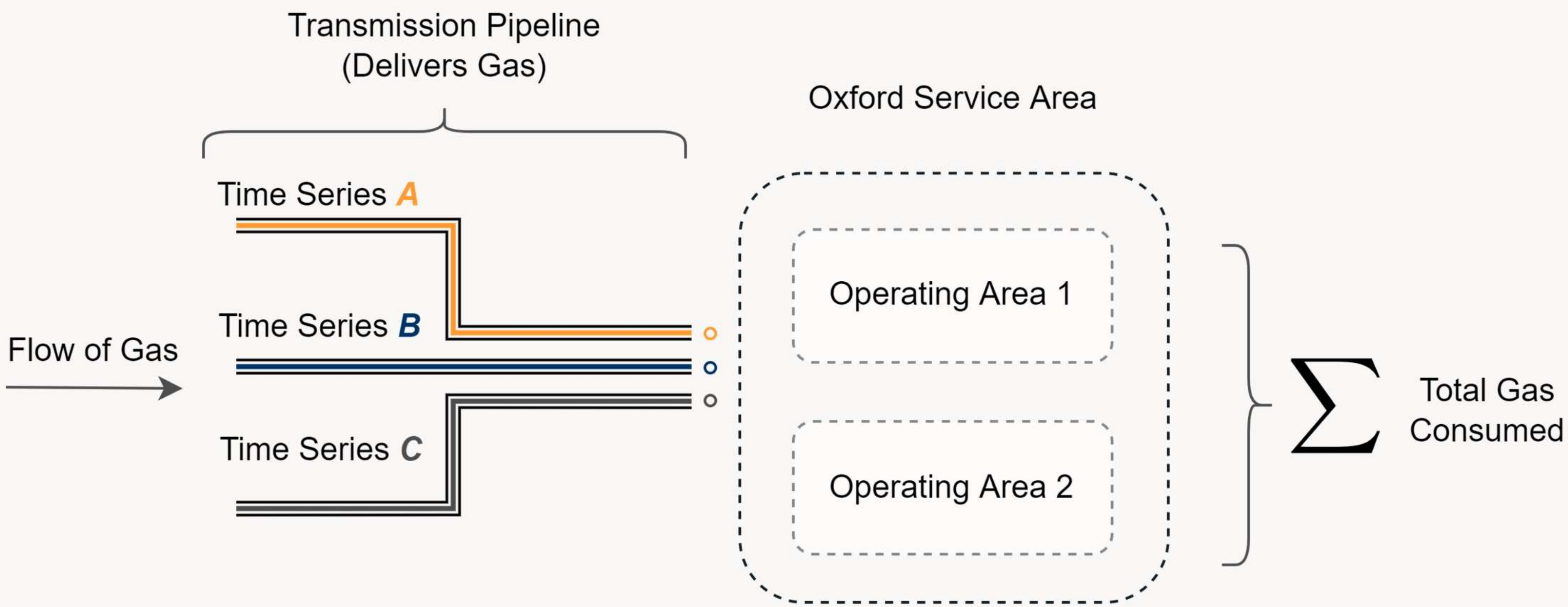
- Natural gas is a fossil fuel energy source extracted for sale and consumption
  - Residential, commercial, and industrial uses
- What happens when gas is not available?
- Gas distribution utilities need to know how much gas is required to adequately service their customers daily

Natural Gas Use in the United Kingdom

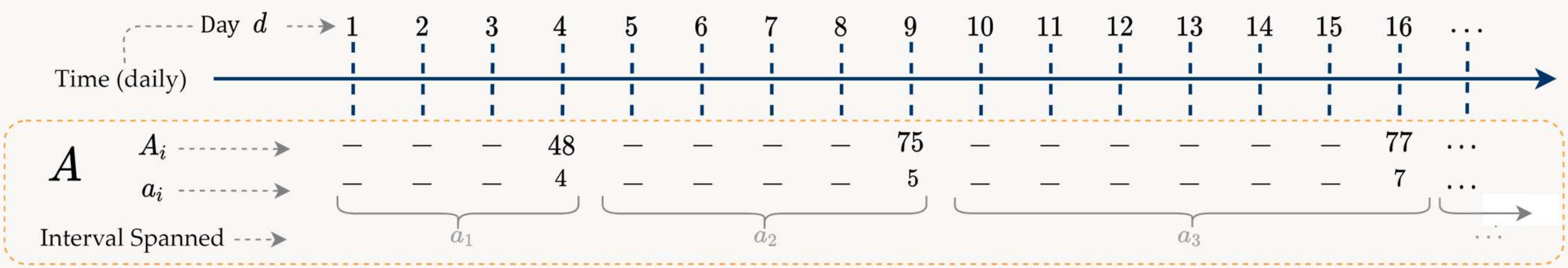
Ranked **11th** in the world for natural gas consumption, accounting for about **2.1%** of the world's total consumption

Coldest date recorded in Oxford, UK: January 14, 1962

Multi-source Iterative Load Shifting Disaggregation Algorithm (MILS)



Temporal Disaggregation



- Utilities want to know how much gas is demanded **daily**, however the existing graduality of A is does not support this
  - The most accurate forecasts are produced when the frequency of measurement matches the frequency demanded

- Disaggregation occurs when quantity A_i is divided into its underlying component parts
  - Example: Interval observation A_2 is disaggregated into its 5 daily components

Temporal Disaggregation

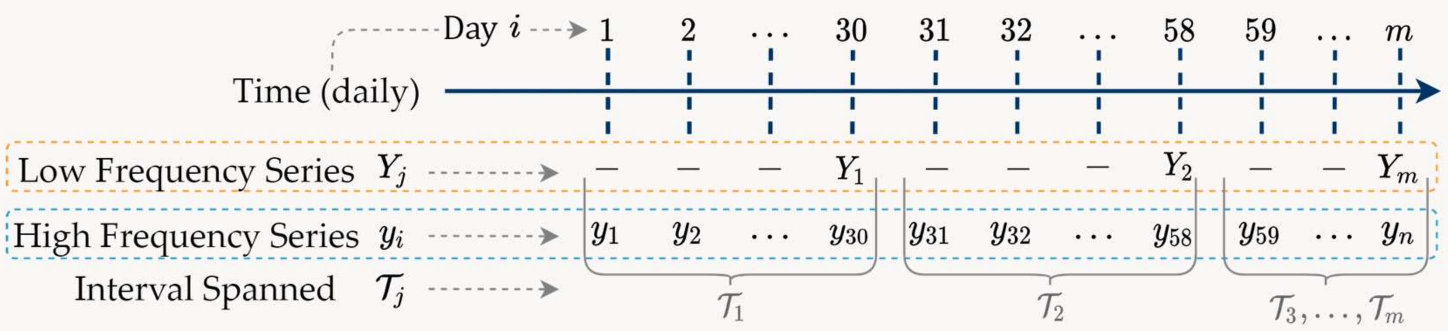
- Given an unequally spaced time series Y whose time steps are relatively long,

$$Y = \{ Y_j, j = 1, \dots, m \},$$

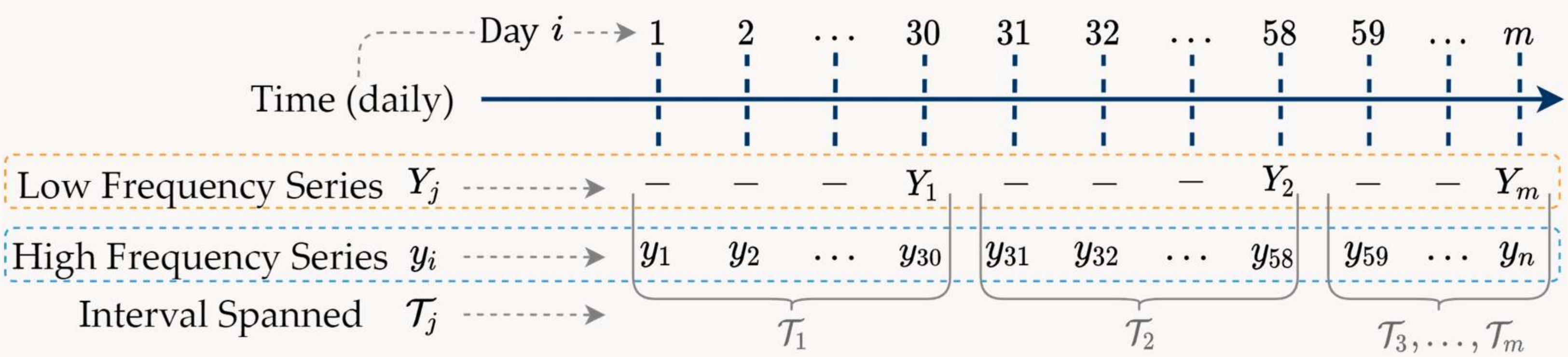
a forecaster may require y

$$y = \{ y_i, i = 1, \dots, n \},$$

an underlying series of higher frequency data for forecasting.



Temporal Disaggregation



- Define aggregation operator $\mathcal{A}$ such that

$$Y = \mathcal{A}(y, \mathcal{T}), \quad \text{where} \qquad Y_j = \sum_{\mathcal{T}_j \in \mathcal{T}} y_i \,.$$

- The inverse of the aggregation operator, $\mathcal{A}^{-1}$, is the disaggregation operator

$$y = \mathcal{A}^{-1}(Y, \mathcal{T})$$

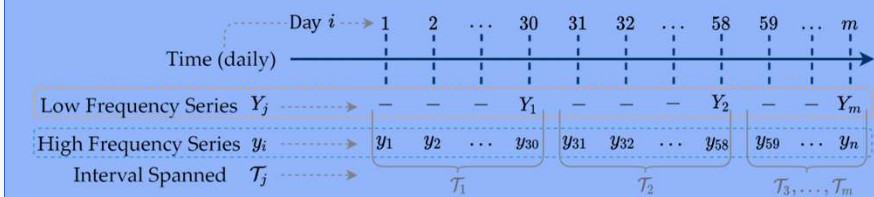
Temporal Disaggregation

- Temporal disaggregation is used to
  - Generate additional, higher frequency historical data
  - Estimate parameters
  - Reintroduce variability into a series that might have been smoothed through aggregation
  - Analyze the data at a resolution previously unavailable

- Current distributive disaggregation techniques
  - Disaggregation problem is ill-posed
  - Working with proportions

Flow Reconciliation Constraint

Disaggregated high frequency values sum to the low frequency interval from which they originate



Minimize $\quad J(y, \hat{y}) \colon e_i = (y_i - \hat{y}_i)^2$

with respect to $\quad \hat{y} \in \mathcal{A}^{-1}(Y, \mathcal{T})$

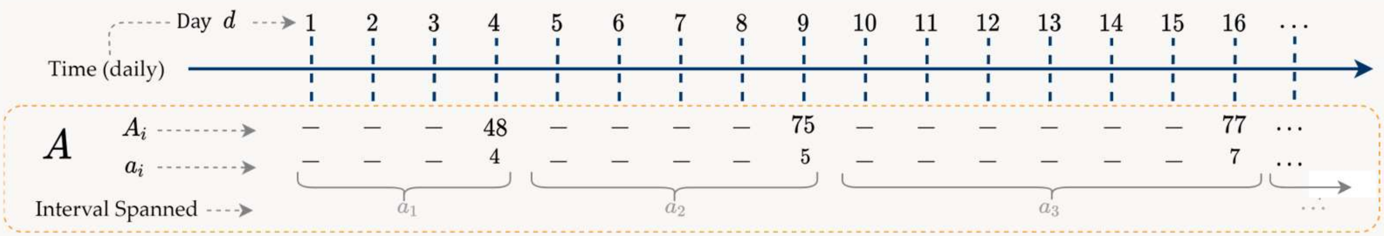
subject to $\quad Y_j = \sum_{\mathcal{T}_j \in \mathcal{T}} \hat{y}_i \quad j = 1, \ldots, m.$

Multi-source Iterative Load Shifting Disaggregation Algorithm (MILS)

- A disaggregation method that accepts multiple data sources structured at nonuniform, low levels of aggregation and outputs a single disaggregated series

- Uses independent variables that are correlated with natural gas consumption to recreate the variability inherent in the underlying series

- Maintains Flow Reconciliation constraint
  - Iterative two-step process
  1. Prediction phase
  2. Update phase



Step 1. Collect MILS inputs

Inputs:

$\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ **–** Multiple low frequency, inconsistently spaced, time series to be disaggregated

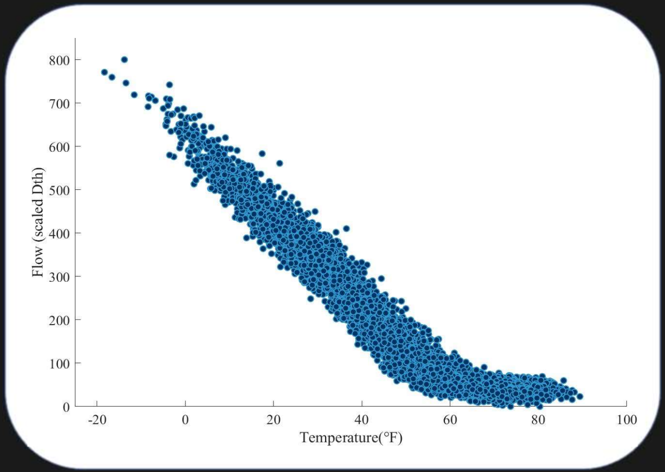
$\boldsymbol{x_s}$ – s independent correlated variables measured at the target frequency of $\widehat{\mathrm{Y}}$

Outputs:

$\widehat{\boldsymbol{Y}}$ – An estimate of daily flow

MILS overview

Step
1. Collect MILS inputs
2. Prepare gas consumption time series inputs
3. Prepare exogenous variable inputs
4. Iteratively model daily flow
5. Shift flow subject to Flow Reconciliation constraint



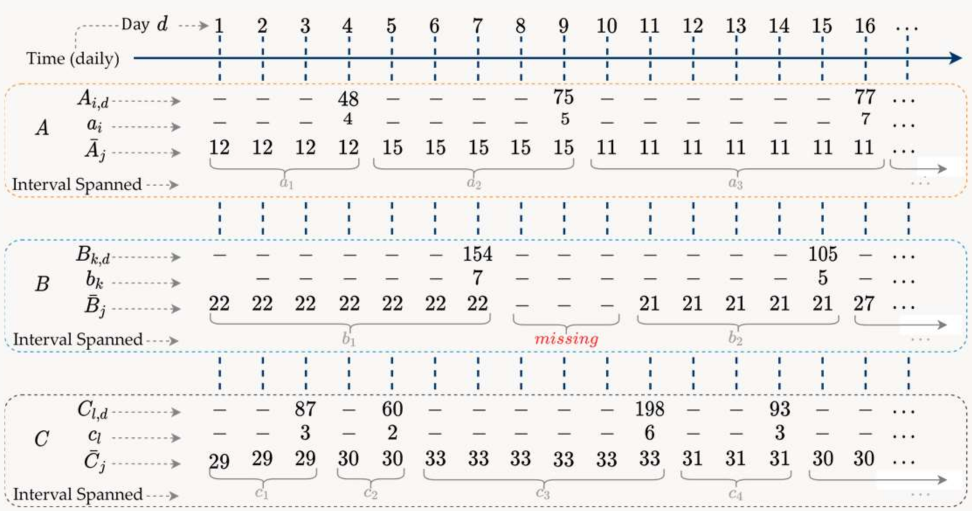
MILS* - Multi-source Iterative Flow Shifting Disaggregation

Step 2. Preparation of gas consumption time series inputs

Given

$$A = \{ A_i, i = 1, \dots, N_A \}$$
$$B = \{ B_k, k = 1, \dots, N_B \}$$
$$C = \{ C_l, l = 1, \dots, N_C \}$$

Naively disaggregate A, B, and C into their daily counterparts $\bar{A}$, $\bar{B}$, and $\bar{C}$



MILS overview

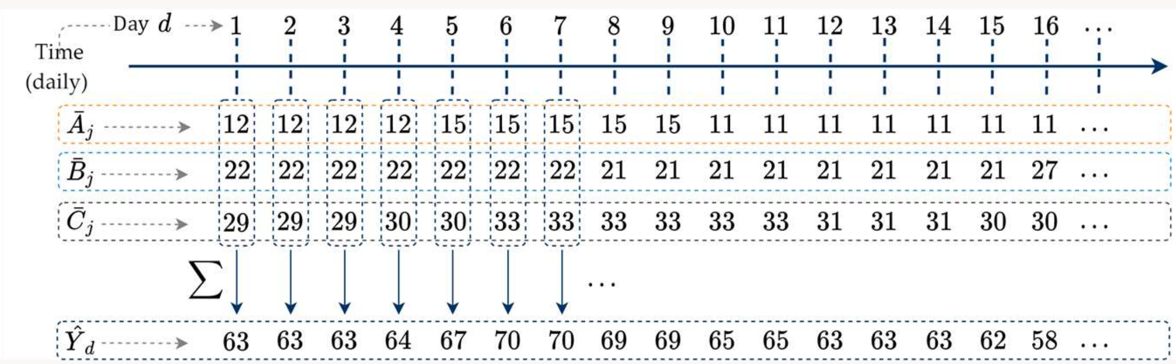
Step
1. Collect MILS inputs
2. Prepare gas consumption time series inputs
3. Prepare exogenous variable inputs
4. Iteratively model daily flow
5. Shift flow subject to Flow Reconciliation constraint

$$\bar{A}_j = \frac{A_i}{a_i}$$

Step 2. Preparation of gas consumption time series inputs

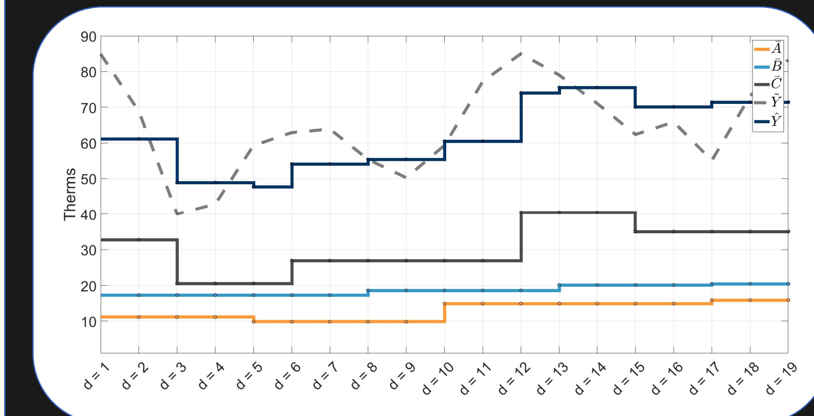
Aggregate $\bar{A}$, $\bar{B}$, and $\bar{C}$ on day d

$$\hat{Y}_d = \bar{A}_d + \bar{B}_d + \bar{C}_d$$



MILS overview

Step
1. Collect MILS inputs
2. Prepare gas consumption time series inputs
3. Prepare exogenous variable inputs
4. Iteratively model daily flow
5. Shift flow subject to Flow Reconciliation constraint



MILS* - Multi-source Iterative Flow Shifting Disaggregation

Step 3. Preparation of exogenous

Select high frequency independent correlated variables proven to be good indicators of future natural gas demand

$$\mathrm{T} = \{ T_d, \mathrm{d} = 1, \dots, N_d\}$$
$$W = \{ W_d, d = 1, \dots, N_d\}$$

Transform to be

$$HDD_d = \max(T_{ref} - T_d, 0),$$

$$HDDW_d = \begin{cases} HDD_{ref} \dfrac{152 + W_d}{160} & W_d \leq 8 \\ HDD_{ref} \dfrac{72 + W_d}{80} & W_d > 8 \end{cases}$$

$$CDD_d = \max(0, T_d - T_{ref})$$

MILS overview

Step
1. Collect MILS inputs
2. Prepare gas consumption time series inputs
3. Prepare exogenous variable inputs
4. Iteratively model daily flow
5. Shift flow subject to Flow Reconciliation constraint

MILS* - Multi-source Iterative Flow Shifting Disaggregation

Step 3. Preparation of exogenous

- Given $x_1, \ \dots, x_s$ independent correlated variables measured daily
  - $x_1 = T_d =$ Temperature
  - $x_2 = W_d =$ Wind

- Apply nonlinear transform
  - $HDD_d = \max(T_{ref} - T_d, 0)$,
  - $HDDW_d = \begin{cases} HDD_{ref} \frac{152 + W_d}{160} & W_d \leq 8 \\ HDD_{ref} \frac{72 + W_d}{80} & W_d > 8 \end{cases}$,
  - $CDD_d = \max(0, T_d - T_{ref})$

- Form design matrix X

MILS overview

Step
1. Collect MILS inputs
2. Prepare gas consumption time series inputs
3. Prepare exogenous variable inputs
4. Iteratively model daily flow
5. Shift flow subject to Flow Reconciliation constraint

$$\mathrm{T} = \{\, T_d, \mathrm{d} = 1, \dots, N_d\}$$
$$W = \{\, W_d, d = 1, \dots, N_d\}$$

$$X = \begin{bmatrix} 1 & HDD_1 & HDDW_1 & CDD_1 \\ 1 & \vdots & \vdots & \vdots \\ \vdots & HDD_{N_D} & HDDW_{N_D} & CDD_{N_D} \end{bmatrix}$$

MILS* - Multi-source Iterative Flow Shifting Disaggregation

Step 4. Iteratively model daily flow

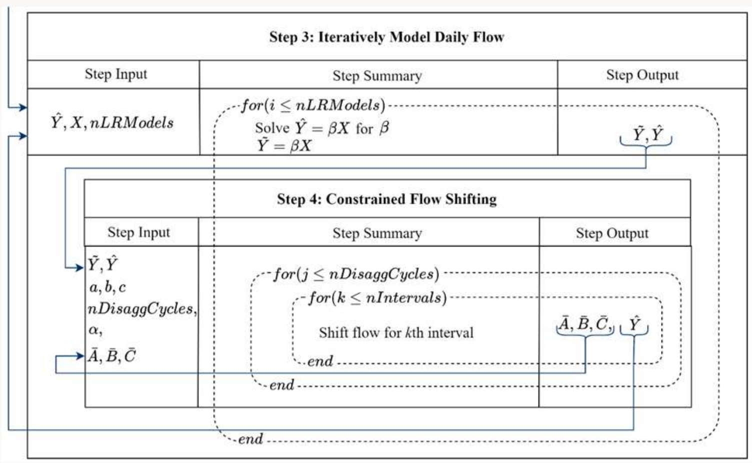
$$\hat{Y} = X\vec{\beta}.$$

Solve for

$$\vec{\beta}.$$

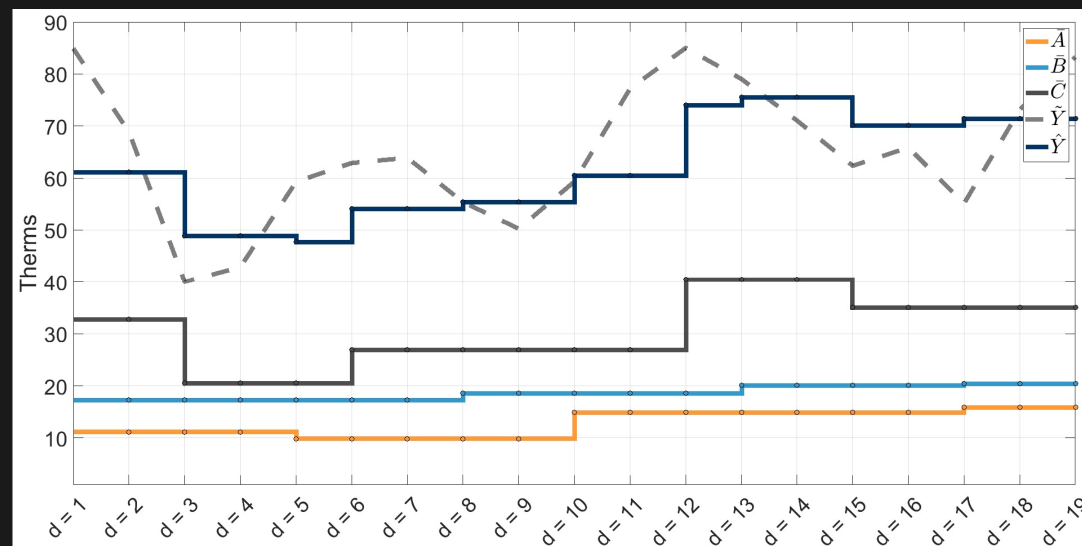
Calculate daily estimates

$$\tilde{Y} = X\vec{\beta}.$$



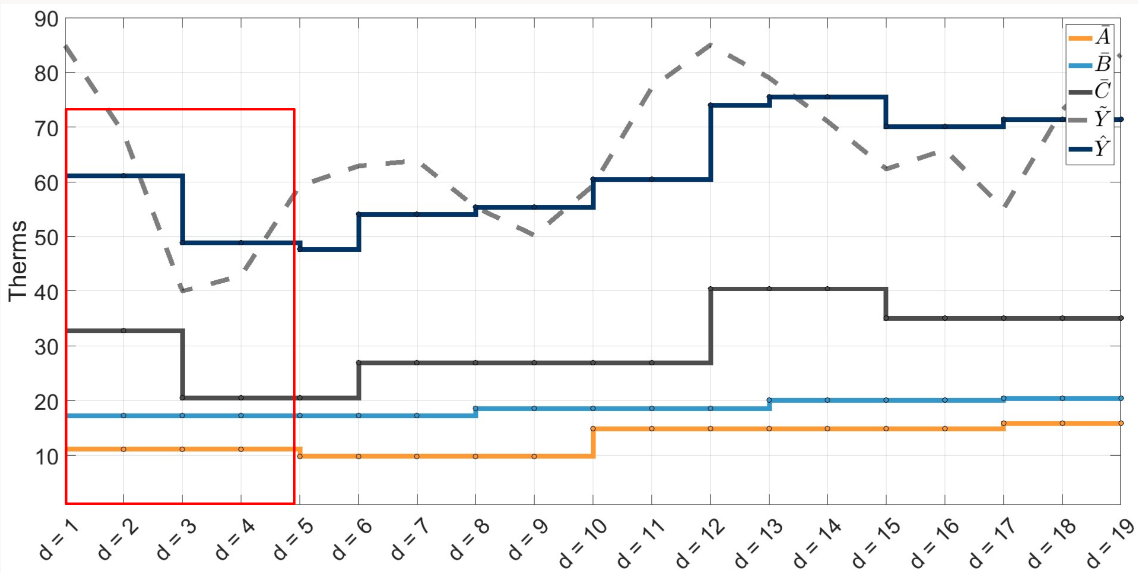
MILS overview

Step
1. Collect MILS inputs
2. Prepare gas consumption time series inputs
3. Prepare exogenous variable inputs
4. Iteratively model daily flow
5. Shift flow subject to Flow Reconciliation constraint



MILS* - Multi-source Iterative Flow Shifting Disaggregation

Step 5. Constrained Flow Shifting

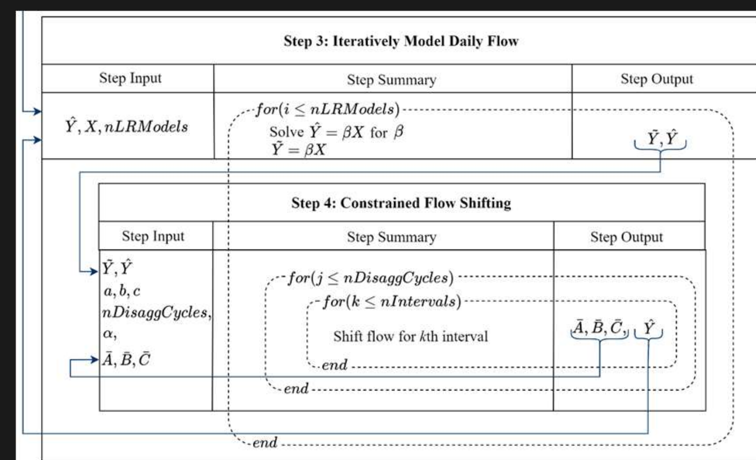


Given
$A_1 = 48$
and
$a_1 = 4$

$\bar{A}_d = 12 \; for \; d = 1, \ldots, 4.$

MILS overview

Step
1. Collect MILS inputs
2. Prepare gas consumption time series inputs
3. Prepare exogenous variable inputs
4. Iteratively model daily flow
5. Shift flow subject to Flow Reconciliation constraint



MILS* - Multi-source Iterative Flow Shifting Disaggregation

Step 5. Constrained Flow Shifting

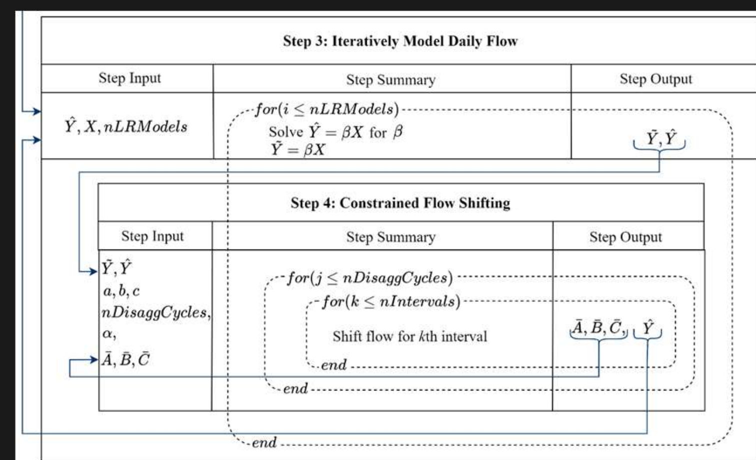
We know A_1 contributes 48 units over days $\hat{Y}_1$, $\hat{Y}_2$, $\hat{Y}_3$, and $\hat{Y}_4$

Remove A_1's contribution to $\hat{Y}$ over days 1-4.

$$\hat{Y_d} = \hat{Y_d} - \bar{A_d} \text{ for } d = 1, 2, \ldots, 4.$$

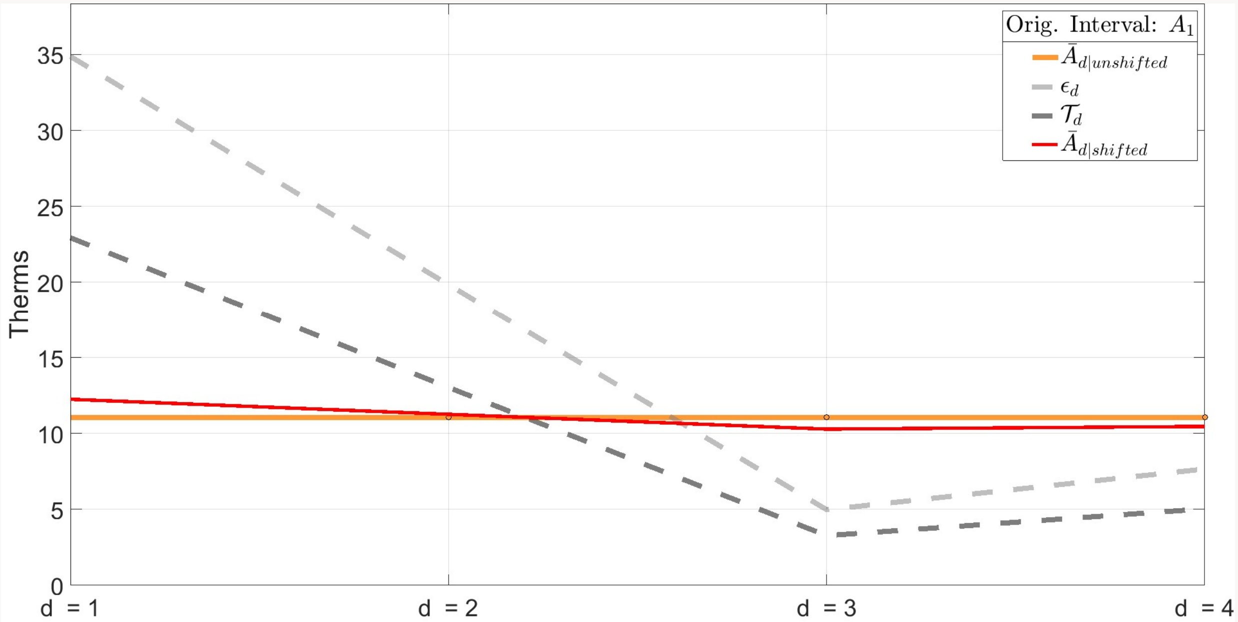
MILS overview

Step
1. Collect MILS inputs
2. Prepare gas consumption time series inputs
3. Prepare exogenous variable inputs
4. Iteratively model daily flow
5. Shift flow subject to Flow Reconciliation constraint



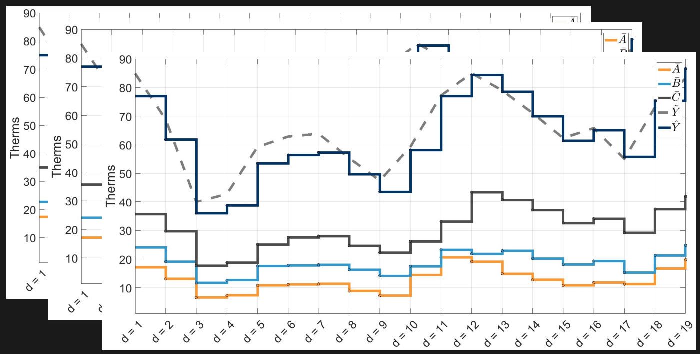
MILS* - Multi-source Iterative Flow Shifting Disaggregation

Step 5. Constrained Flow Shifting



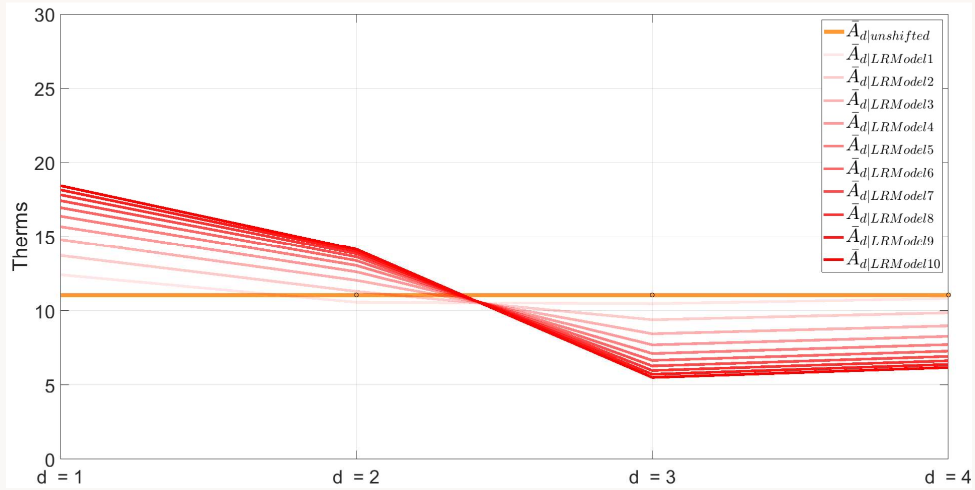
MILS overview

Step
1. Collect MILS inputs
2. Prepare gas consumption time series inputs
3. Prepare exogenous variable inputs
4. Iteratively model daily flow
5. Shift flow subject to Flow Reconciliation constraint



MILS* - Multi-source Iterative Flow Shifting Disaggregation

Case Study Results



- Multi-source Iterative Least Squares (MILS) Performance
  - Three year period
- Benchmark models:
  - Naïve Disaggregation (NAÏVE) [3], Generalized Least Squares (GLS) [18], ARIMA [26]
  - Error Metrics:
    - RMSE - Dth
    - MAPE - % error

	Disaggregation Method			
Metric	MILS	NAIVE	GLS	ARIMA
RMSE	106.12 Dth	380.56 Dth	213.72 Dth	154.12 Dth
MAPE	8.60%	30.87%	17.32%	12.49%

Questions?

Marquette
Energy
Analytics

Thank you.

info@marquetteenergyanalytics.com
(414) 765-2839

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

Appendix

[1] B. Gal, “The Challenges of Time-Series Forecasting in Retail,” *Retail Technology*, 2017. https://cb4.com/blog/time-series-forecasting-challenges/ (accessed Mar. 20, 2021).
[2] S. R. Vitullo, G. F. Corliss, M. Adya, F. Nourzad, and R. H. Brown, “An algorithm for disaggregating temporal natural gas consumption,” *Can. Appl. Math. Q.*, vol. 21, no. 3, pp. 391–410, 2013.
[3] W. Wei and D. Stram, “Disaggregation of Time Series Models,” *Wiley R. Stat. Soc.*, vol. 52, no. 3, pp. 453–467, 1990, [Online]. Available: https://www.jstor.org/stable/2345669.
[4] F. Moauro and G. Savio, “Temporal disaggregation using multivariate structural time series models,” *Econom. J.*, vol. 8, no. 2, pp. 214–234, 2005, doi: 10.1111/j.1368-423x.2005.00161.x.
[5] A. Dotis-Georgiou, “Autocorrelation in Time Series Data,” *DZone: A Devada Media Property*, 2019. https://dzone.com/articles/autocorrelation-in-time-series-data (accessed Jan. 02, 2020).
[6] H. Akaike, “Fitting Autoregressive Models for Prediction,” *Ann. Inst. Stat. Math.*, vol. 21, no. 1, pp. 243–247, 1969, doi: 10.1007/BF02532251.
[7] J. Durbin, “The Fitting of Time-Series Models,” *Int. Stat. Inst.*, vol. 28, no. 3, pp. 233–244, 1960, [Online]. Available: https://www.jstor.org/stable/1401322.
[8] V. Flovik, “How (not) to use Machine Learning for time series forecasting: Avoiding the pitfalls,” *Towards Data Science*, 2018. https://towardsdatascience.com/how-not-to-use-machine-learning-for-time-series-forecasting-avoiding-the-pitfalls-19f9d7adf424 (accessed Jan. 07, 2020).
[9] A. Harvey, “Forecasting with Unobserved Components Time Series Models,” in *Handbook of Economic Forecasting*, vol. 1, G. Elliott, C. Granger, and A. Timmermann, Eds. Elsevier, 2006, pp. 327–412.
[10] C. Chafield, *Time-Series Forecasting*, 1st ed., no. 1. Boca Raton, Florida: Chapman & Hall, 2000.
[11] “Time Series Analysis: Concept, Additive and Multiplicative Models,” *Theintactone*, 2020. https://theintactone.com/2019/08/28/bsa-u2-topic-1-time-series-analysis-concept-additive-and-multiplicative-models/ (accessed Apr. 02, 2021).
[12] T. Hong, “Energy Forecasting: Past, Present, and Future,” *Foresight Int. J. Appl. Forecast.*, vol. 32, pp. 43–48, 2013.
[13] B. Soldo, “Forecasting natural gas consumption,” *Appl. Energy*, vol. 92, pp. 26–37, 2012, doi: 10.1016/j.apenergy.2011.11.003.
[14] S. Makridakis and R. L. Winkler, “Averages of Forecasts: Some Empirical Results,” *Manage. Sci.*, vol. 29, no. 9, pp. 987–996, Sep. 1983, doi: 10.1287/mnsc.29.9.987.
[15] B. E. Flores, “A pragmatic view of accuracy measurement in forecasting,” *Omega*, vol. 14, no. 2, pp. 93–98, 1986, doi: https://doi.org/10.1016/0305-0483(86)90013-7.
[16] R. Shumway and D. Stoffer, *Time Series Analysis and Its Applications With R Examples*, 4th ed. Springer, 2017.
[17] P. Balestra and M. Nerlove, “Pooling Cross Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas,” *Econometrica*, vol. 34, no. 3, pp. 585–612, 1966, [Online]. Available: https://www.jstor.org/stable/1909771.
[18] C. Chow and A. Lin, “Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series,” *Rev. Econ. Stat.*, vol. 53, no. 4, pp. 372–375, 1971.
[19] S. H. Mohr and G. M. Evans, “Long term forecasting of natural gas production,” *Energy Policy*, vol. 39, no. 9, pp. 5550–5560, 2011, doi: 10.1016/j.enpol.2011.04.066.
[20] S. R. Vitullo, R. H. Brown, G. F. Corliss, and B. M. Marx, “Mathematical Models for Natural Gas Forecasting,” *Can. Appl. Math Q.*, vol. 17, no. 4, pp. 1–13, 2010, [Online]. Available: papers2://publication/uuid/5E8CDA33-908F-4B19-8F90-02A304882841.
[21] S. Fan and R. J. Hyndman, “Short-term load forecasting based on a semi-parametric additive model,” *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 134–141, 2012, doi: 10.1109/TPWRS.2011.2162082.
[22] E. Elattar, J. Y. Goulermas, and Q. H. Wu, “Electric Load Forecasting Based on Locally Weighted Support Vector Regression,” *IEEE Trans. Syst. Man. Cybern.*, vol. 40, no. 4, pp. 438–447, 2010.
[23] W. S. Cleveland and S. J. Devlin, “Locally weighted regression: An approach to regression analysis by local fitting,” *J. Am. Stat. Assoc.*, vol. 83, no. 403, pp. 596–610, 1988, doi: 10.1080/01621459.1988.10478639.
[24] J. G. Tamba *et al.*, “Forecasting natural gas: A literature survey,” *Int. J. Energy Econ. Policy*, vol. 8, no. 3, pp. 216–249, 2018.
[25] I. Moghram and S. Rahman, “Analysis and Evaluation of Five Short-Term Load Forecasting Techniques,” *IEEE Power Eng. Rev.*, vol. 9, no. 11, pp. 42–43, 1989, doi: 10.1109/MPER.1989.4310383.
[26] C. N. Babu and B. E. Reddy, “A Moving-average Filter Based Hybrid ARIMA-ANN Model for Forecasting Time Series Data,” *Appl. Soft Comput. J.*, vol. 23, pp. 27–38, 2014, doi: 10.1016/j.asoc.2014.05.028.
[27] J. W. Taylor, “Short-term electricity demand forecasting using double seasonal exponential smoothing,” *J. Oper. Res. Soc.*, vol. 54, no. 8, pp. 799–805, 2003, doi: 10.1057/palgrave.jors.2601589.
[28] S. J. Huang and K. R. Shih, “Short-term load forecasting via ARMA model identification including non-Gaussian process considerations,” *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 673–679, 2003, doi: 10.1109/TPWRS.2003.811010.
[29] H. Spliid, “A Fast Estimation Method for the Vector Autoregressive Moving Average Model with Exogenous Variables,” *J. Am. Stat. Assoc.*, vol. 78, no. 384, pp. 843–849, Dec. 1983, doi: 10.1080/01621459.1983.10477030.