COLLABORATIVE INFERENCE OF MISSING SMART ELECTRIC METER DATA FOR A BUILDING

2019 IEEE INTERNATIONAL WORKSHOP ON MACHINE LEARNING FOR SIGNAL PROCESSING

Nan Duan, Jose Cadena, Pedro Sotorrio, Jhi-Young Joo
Lawrence Livermore National Laboratory

October 15, 2019

LLNL-PRES-XXXXXX
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

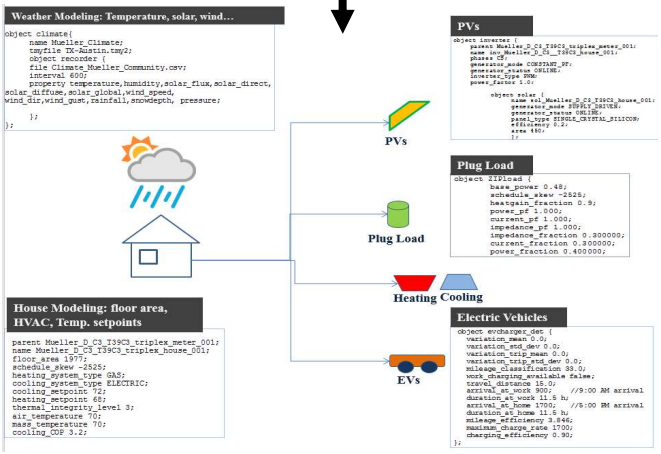
Lawrence Livermore National Laboratory

Background


Smart Meter

opportunity and challenges for modern power grids:

- increasing penetration of renewable energy resources
- flexible/controllable load
- data availability
- customer privacy and cyber-security
- model complexity and accuracy
- transition from deterministic to stochastic planning and operation

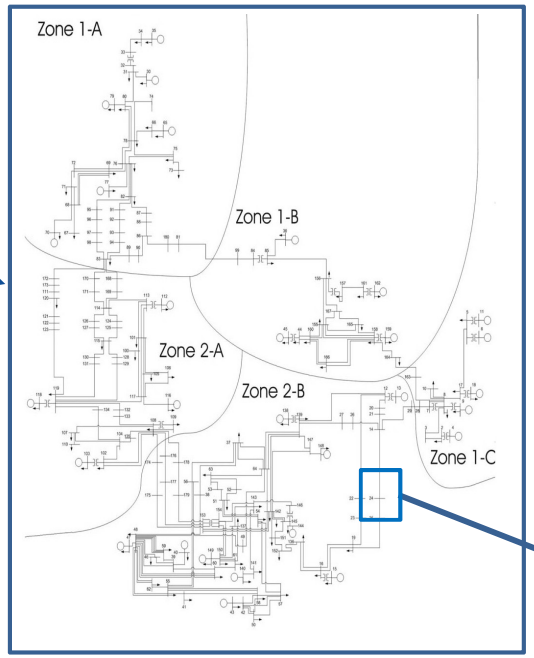


http://gridlab-d.shoutwiki.com/wiki/GridLAB-D_Wiki:GridLAB-D_Tutorial_Chapter_3_-_Basic_Distribution_System_Modeling

Motivation

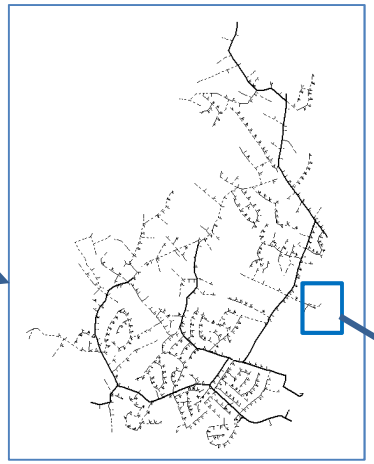
- Model becomes **more uncertain** and **less accurate** moving from balancing authority level to customer level.
- **Smart meter data** can be used to build surrogate model at customer level.

Balancing Authorities

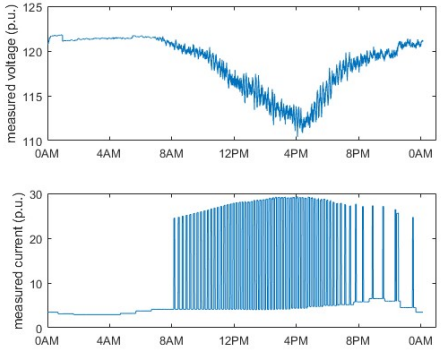


Transmission

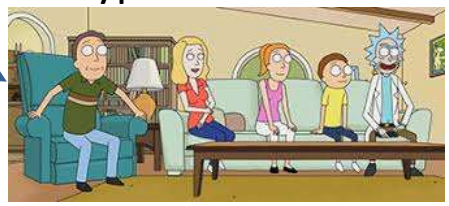
Distribution



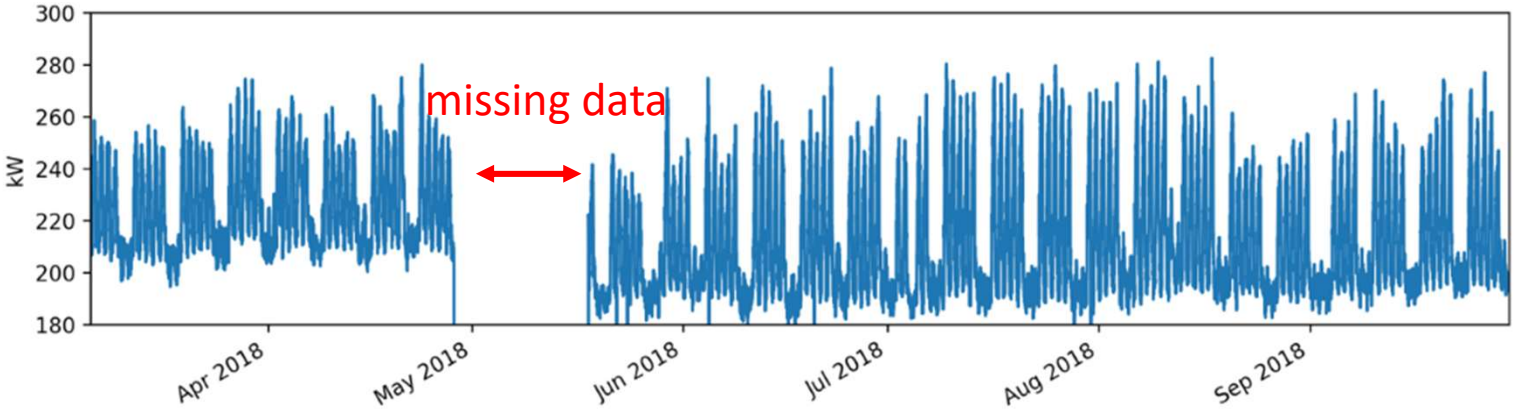
Smart Meter Data



A Typical Household



Motivation

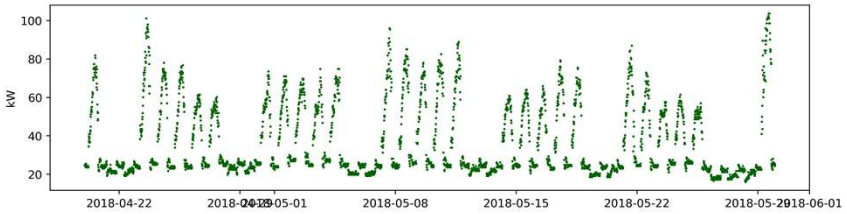


requirements for a surrogate model:

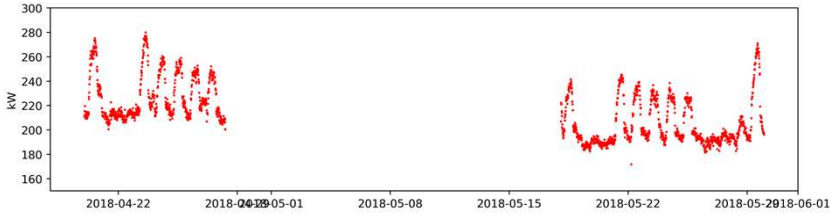
- It needs to be computationally cheap, due to the **limited** processing power of a smart meter
- It needs to capture individual demands or renewable generation resources that are highly influenced by rapid **exogenous** random events, such as weather conditions or individual human behavior.

Formulation

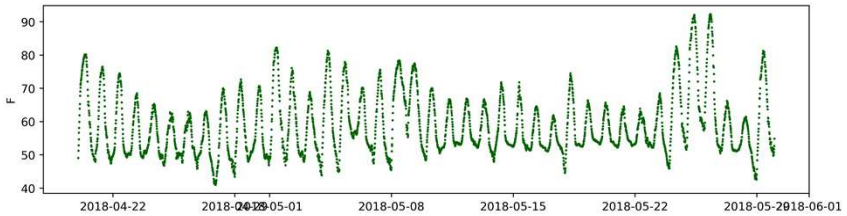
Support Building Meter Data (exogenous)



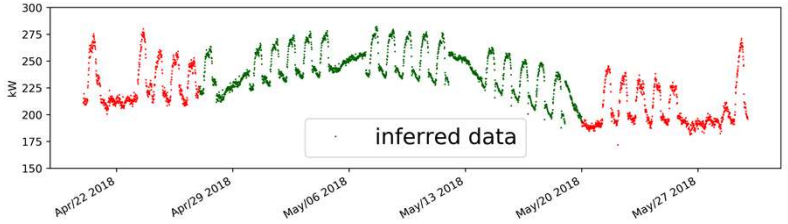
Building with Missing Meter Data



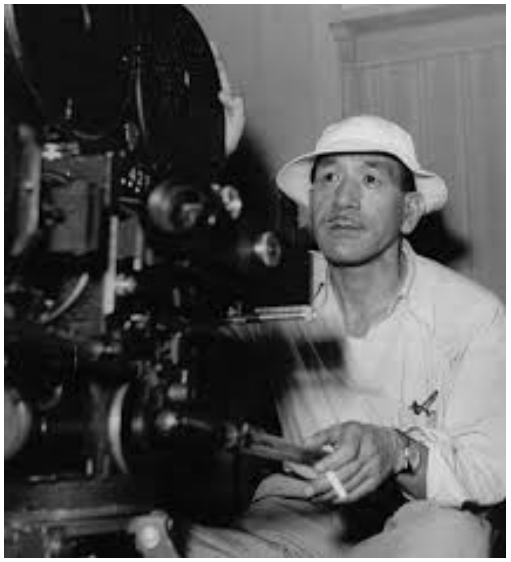
Temperature Data (exogenous)



Surrogate model

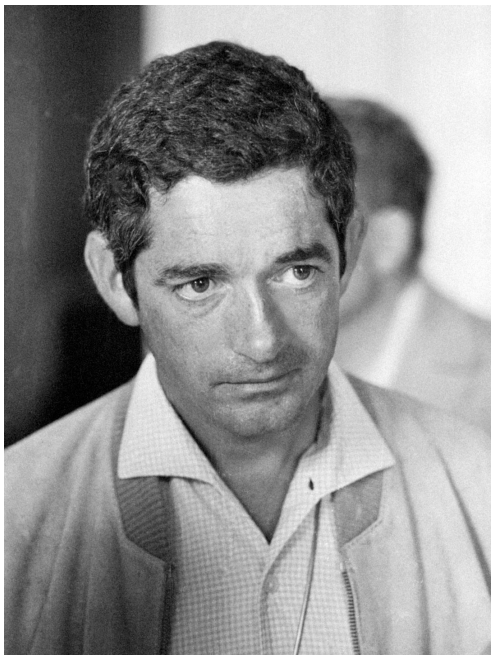


Formulation (An Analogy)



Yasujirō Ozu
Japanese Director,
invented “tatami shot”

Any connection
between them?



Jacques Demy
French New Wave Director,
his movies inspired *La La Land*

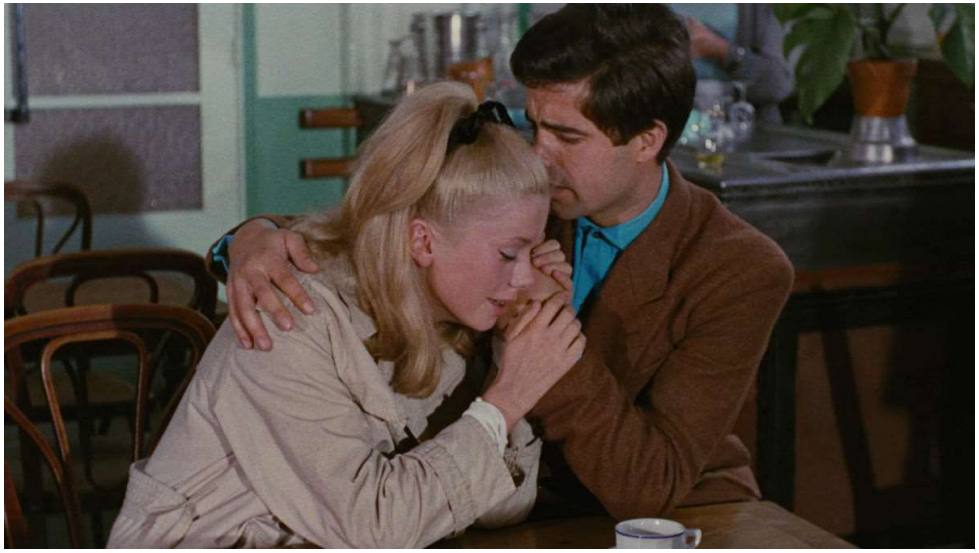
https://www.theguardian.com/film/yasujiro-ozu
https://www.goldenglobes.com/person/jacques-demy

Formulation (An Analogy)

Film Color Faded

Equinox Flower, 1958

Similar Color Palette

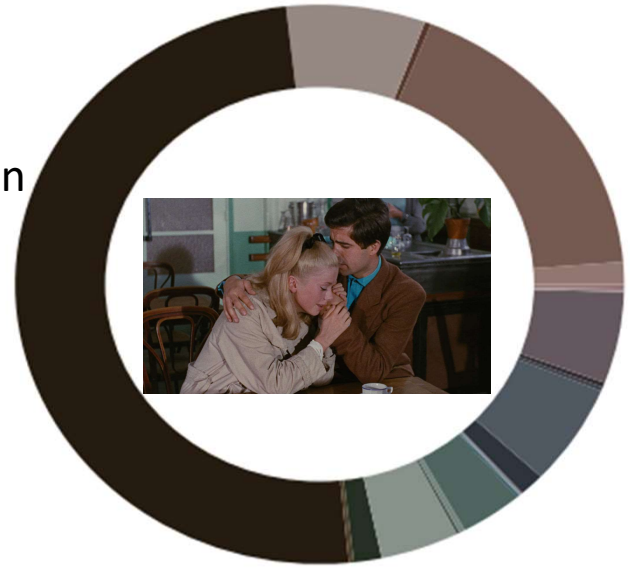


The Umbrellas of Cherbourg, 1964

http://www.filmsufi.com/2014/08/equinox-flower-yasujiro-ozu-1958.html
https://www.criterion.com/current/posts/3235-the-umbrellas-of-cherbourg-a-finite-forever

Formulation (An Analogy)

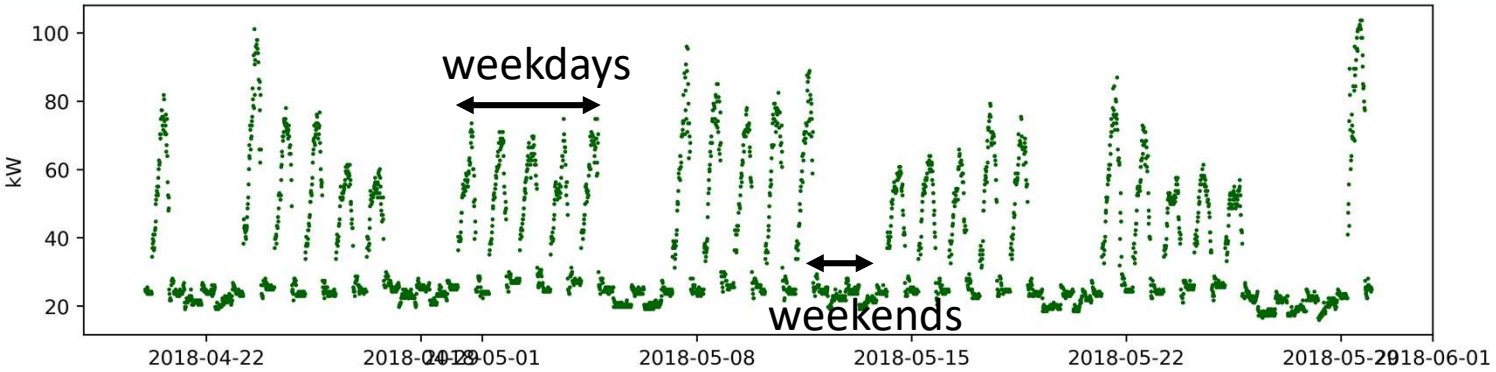
Color Restoration

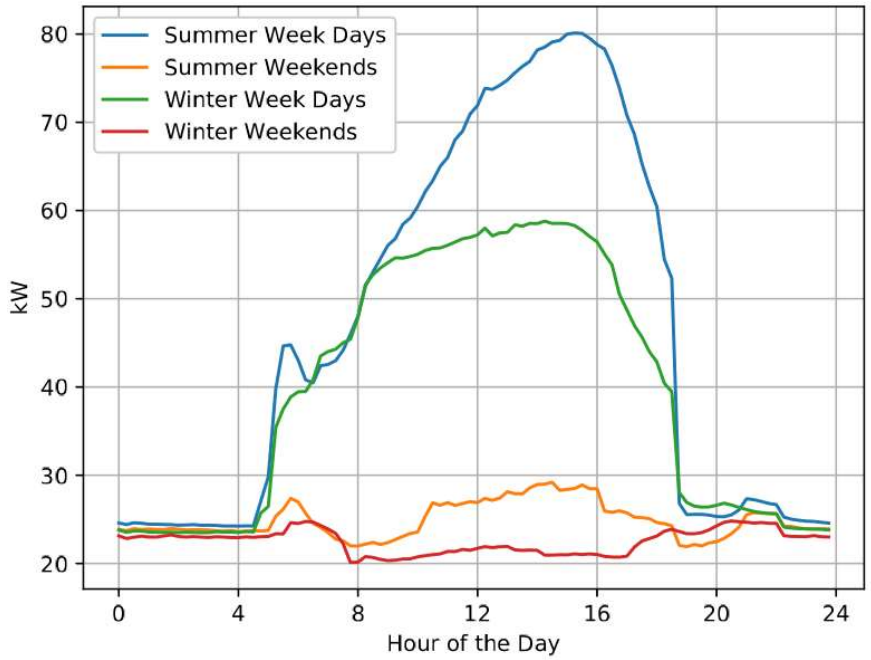


Equinox Flower, 1958

The Umbrellas of Cherbourg, 1964

Data Preparation



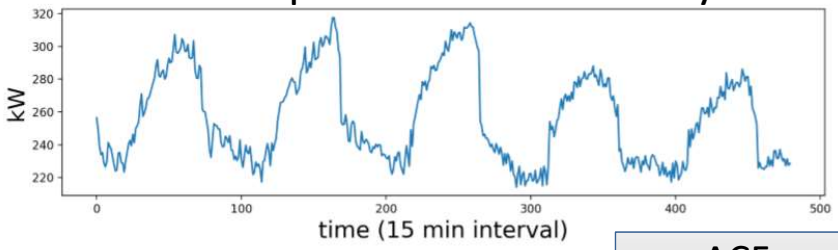


8 datasets for each building

Spring	• Spring Weekdays • Spring Weekends
Summer	• Summer Weekdays • Summer Weekends
Fall	• Fall Weekdays • Fall Weekends
Winter	• Winter Weekdays • Winter Weekends

Data Analysis

Example: summer weekdays



$$m = 24hr \times \frac{60min/hr}{15min}$$

ACF calculation

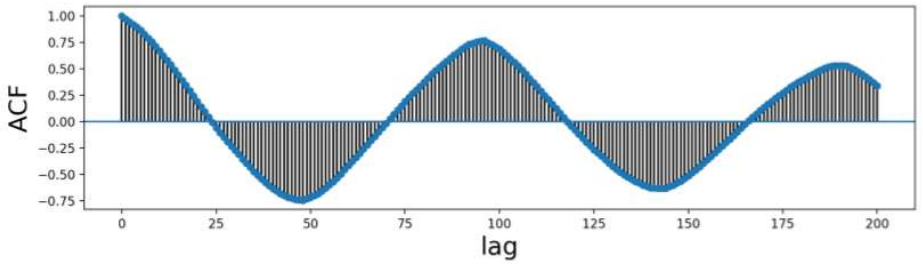
seasonal differentiation

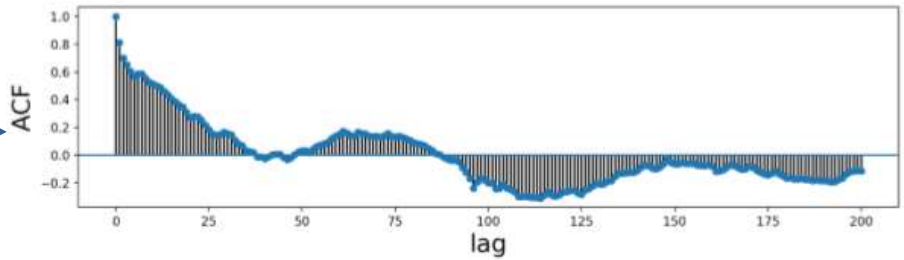
ACF calculation

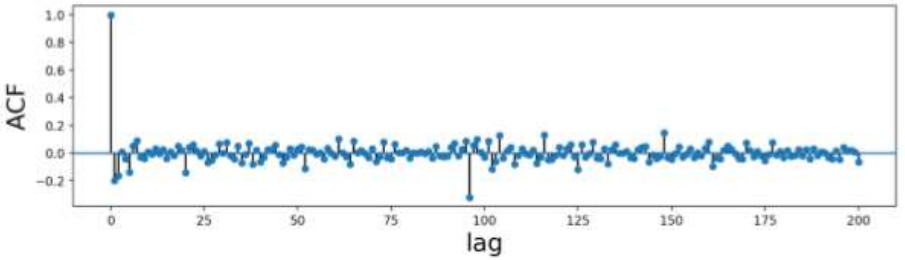
1st order differentiation

ACF calculation

auto-correlation function (ACF)

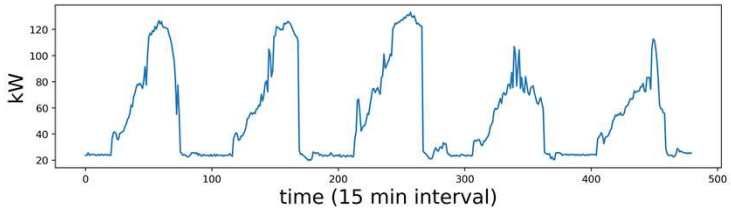






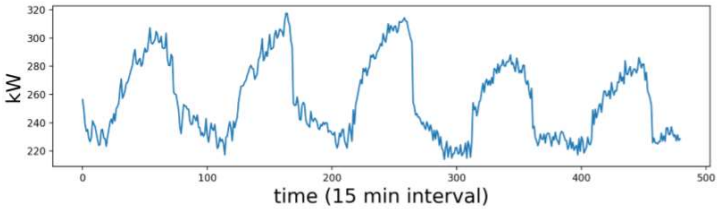
Data Analysis

Support Building Meter Data (exogenous)



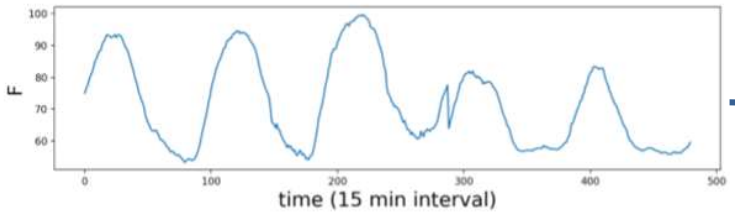
seasonal and 1st order differentiation

Building to be Modeled



seasonal and 1st order differentiation

Temperature Data (exogenous)

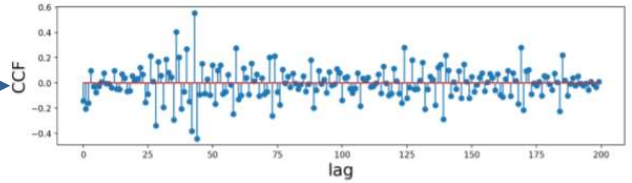


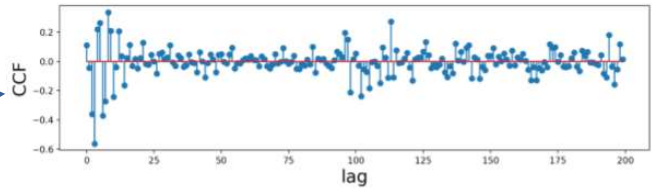
seasonal and 1st order differentiation

CCF calculation

CCF calculation

cross-correlation function (CCF)





Surrogate Model

seasonal autoregressive integrated moving average model with exogenous variables (SARIMAX)

$$\Phi(B^P)\phi(B^p)\nabla_m^D\nabla_1^d Y_t = \Theta(B^Q)\theta(B^q)Z_t$$

$$B^n Y_t = Y_{t-n}$$

$$\Phi(B^P) = 1 - \Phi_1 B^m - \ldots - \Phi_P B^{Pm}$$

$$\phi(B^p) = 1 - \phi_1 B - \ldots - \phi_P B^p$$

$$\Theta(B^Q) = 1 + \Theta_1 B^m + \ldots + \Theta_Q B^{Qm}$$

$$\theta(B^q) = 1 + \theta_1 B + \ldots + \theta_P B^q$$

$$\nabla_m Y_t = Y_t - Y_{t-m}$$

$$Z_t \sim N(0, \sigma_Z^2)$$

inclusion of exogenous variable(s):

$$\nabla_m^D\nabla_1^d Y_t = C + v(B^e)\nabla_m^D\nabla_1^d X_t + \frac{\Theta(B^Q)\theta(B^q)}{\Phi(B^P)\phi(B^p)}Z_t$$

$$v(B^e) = v_0 + v_1 B + v_2 B^2 + \ldots + v_e B^e$$

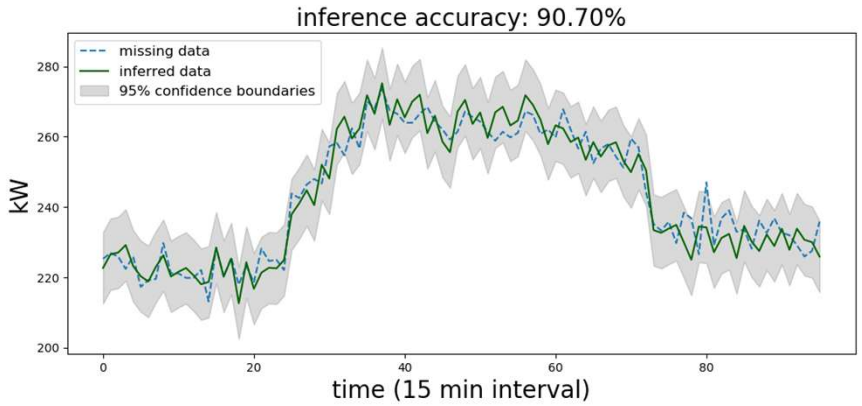
Meter Data Inference Procedure:

1. Data preprocessing and segregation.
2. Parameter setting.
3. Leveraging observed time series data.
4. Model training.
5. Filling the gaps.

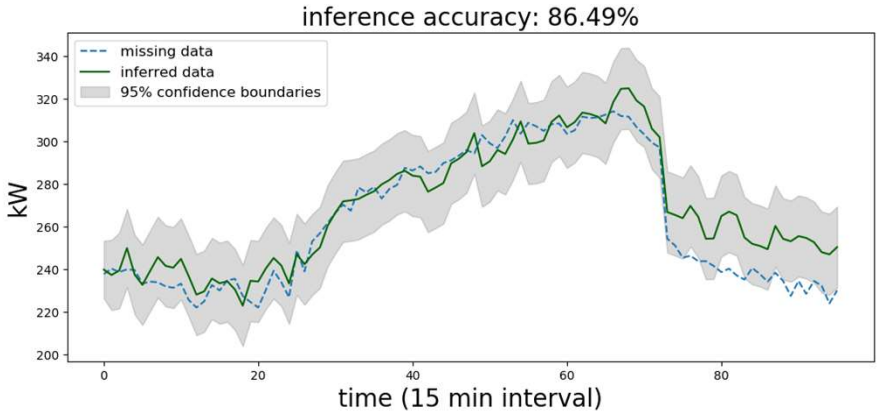
Python package pmdarim is used for model training and testing.
https://pypi.org/project/pmdarima/

Results (Weekdays)

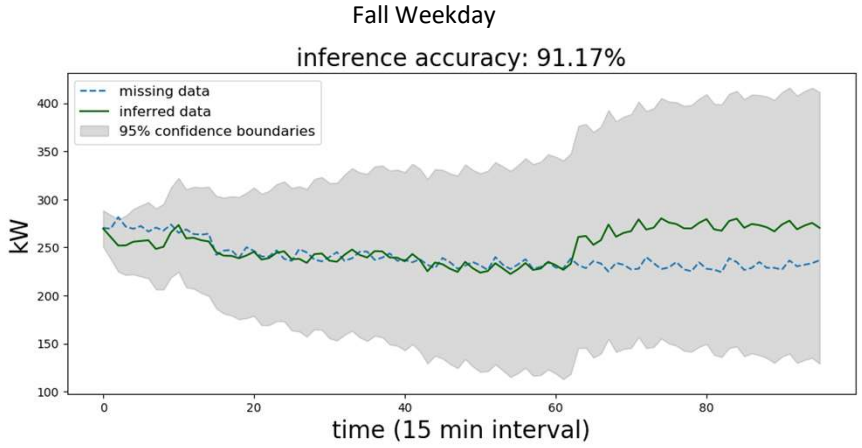
Spring Weekday



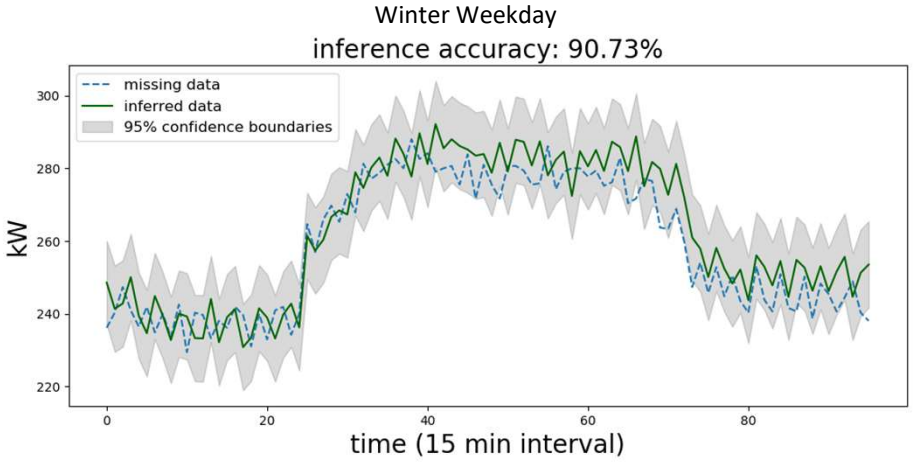
Summer Weekday



Fall Weekday



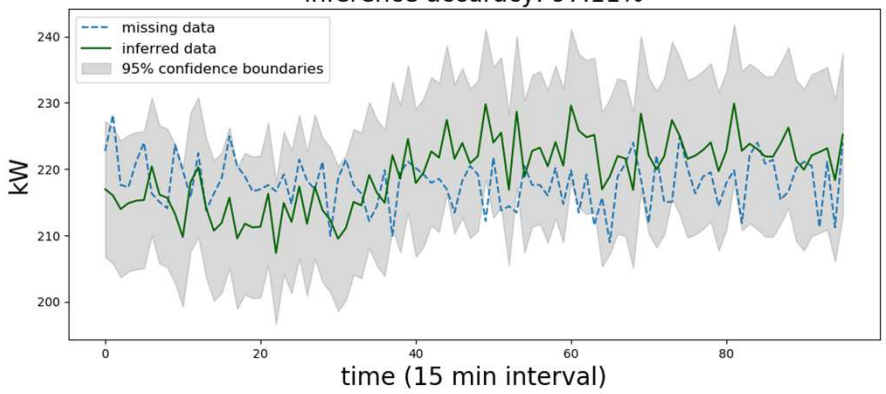
Winter Weekday



Results (Weekends)

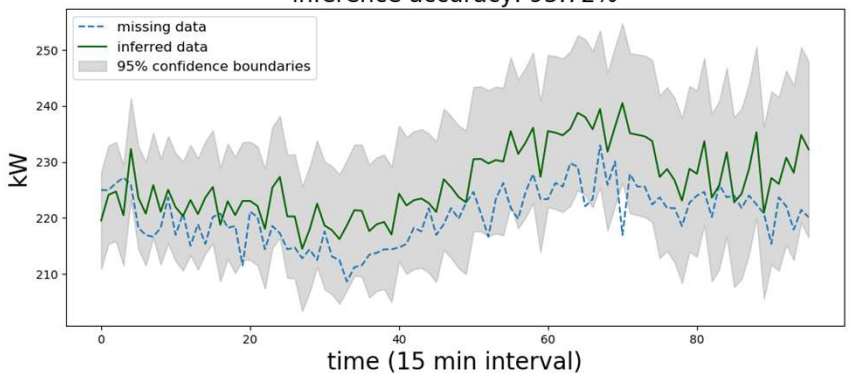
Spring weekend

inference accuracy: 97.11%



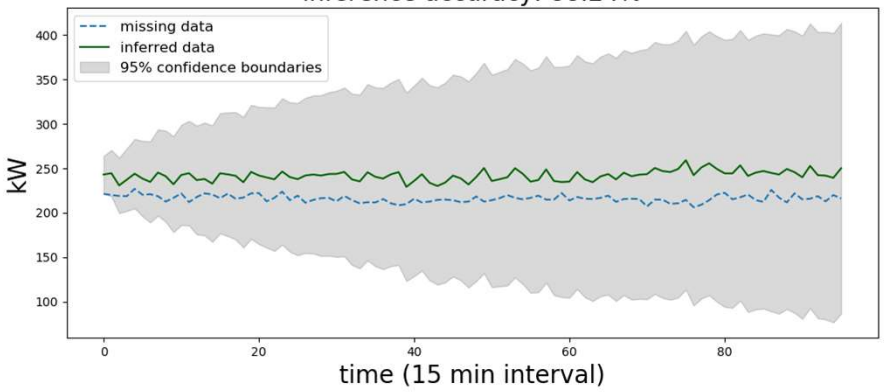
Summer Weekend

inference accuracy: 95.72%



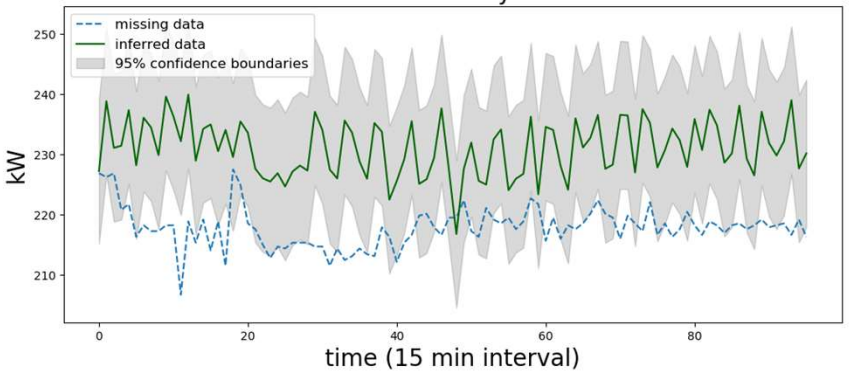
Fall Weekend

inference accuracy: 88.24%



Winter Weekend

inference accuracy: 93.67%



Conclusion and Future Work

- A total of 8 SARIMAX models are trained to infer missing electricity usage data of a building in Martinez, California for weekdays and weekends of different seasons
- The over all inference accuracy is over 85% for all cases.
- For future work, we will investigate the benefit of considering more exogenous variables such as the target building's occupancy and more than one supporting buildings' electricity usage.

Acknowledgement:
This work is partly supported by the Office of Energy Efficiency and Renewable Energy of the U.S. Department of Energy with IM release number LLNL-CONF-77942.