

CONCLUSION

The goal of this research is to apply high dimensional big data analysis method to calibrate building energy model. To achieve the goal, this paper briefly discussed background and theory of LASSO and SCAD method. Compared with conventional variable selection approaches, our methods are robust to deal with insufficient data observations, noise and outliers. More importantly, these two methods help us prioritize calibration effort. In the research, we proposed a 3-step calibration strategy and explained the idea and method for each step. Given limited staff resource, time and budget in the actual project, this strategy significantly reduces the time and effort on calibration and achieve high model accuracy. A 2 million square feet office building campus data in the U.S northeast region is used as case study to verify the method in application.

The advantages of this approach over traditional methods are discussed. The calibrated model can be used to analyze savings for different ECMs and applied in measurement and verification. Selected variables from LASSO and SCAD method are key variables that relate to the building energy use. In the energy management system, those variables should be monitored continuously and compared with historical data. This will show the energy flow and trends inside the building. This calibration steps can be followed in other building types and data.

Reference

- Ahmad, M. and Culp, C.H., 2006. Uncalibrated building energy simulation modeling results. HVAC&R Research, 12(4), pp.1141-1155.
- Department of Energy (DOE). 2011. Building Energy Data Book.
- Energy Information Administration (EIA). 2017. Commercial Building Energy Consumption Survey (CBECS)
- Fan, J. and Li, R., 2001. Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of the American statistical Association, 96(456), pp.1348-1360.
- Fan, J. and Li, R., 2006. Statistical challenges with high dimensionality: Feature selection in knowledge discovery. arXiv preprint math/0602133.
- O'Donnell, J.T., 2009. Specification of optimum holistic building environmental and energy performance information to support informed decision making.
- Raftery, P., Keane, M. and Costa, A., 2009, July. Calibration of a detailed simulation model to energy monitoring system data: a methodology and case study. In Eleventh International IBPSA Conference (pp. 1199-1206).
- Song, S. and Haberl, J., 2008. A procedure for the performance evaluation of a new commercial building: Part I-Calibrated as-built simulation.
- Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological), pp.267-288.
- Wang, L., Kim, Y. and Li, R., 2013. Calibrating non-convex penalized regression in ultra-high dimension. Annals of statistics, 41(5), p.2505.
- Westphal, F.S. and Lamberts, R., 2005, August. Building simulation calibration using sensitivity analysis. In Ninth International IBPSA Conference (pp. 1331-1338).
- Zou, H. and Li, R., 2008. One-step sparse estimates in nonconcave penalized likelihood models. Annals of statistics, 36(4), p.1509.

Nomenclature

- Y observations of response variable (total energy use).
- X observations of independent variables.
- β the coefficients or parameters of linear regression model (1).
- $\hat{\beta}$ the statistical estimator by using the observations.
- λ the tuning parameter for LASSO or SCAD method.
- n the sample size (observation days)
- p the total number of considered independent variables
- P_λ the penalty function with tuning parameter λ .
- ε random error terms.
- u a p-dimensional vector.
- R^p p-dimensional real space.
- $|\cdot|$ the absolute value.
- $\|\cdot\|_2$ the Euclidean norm.

Subscripts

- i the i-th days
- j the j-th considered independent variable.