Towards IoT-Driven Indoor Wellbeing Optimization and Control Using Multivariate Analysis

Published Version:
Afif, O., Ingenito, G., Tartagni, M., Romani, A. (2025). Towards IoT-Driven Indoor Wellbeing Optimization and Control Using Multivariate Analysis. IEEE [10.1109/IWASI66786.2025.11121962].

Availability:
This version is available at: https://hdl.handle.net/11585/1039384 since: 2026-01-26

Published:
DOI: http://doi.org/10.1109/IWASI66786.2025.11121962

Towards IoT-Driven Indoor Wellbeing Optimization and Control Using Multivariate Analysis

Oumaima Afif, Gaetano Ingenito, Marco Tartagni, and Aldo Romani

e-mail: {oumaima.afif2, marco.tartagni, aldo.romani}@unibo.it, {gaetano.ingenito}@studio.unibo.it

***Abstract*—Wellness and comfort are key to occupant health, productivity, and satisfaction. The rise of Internet of Things (IoT) is transforming building management by integrating advanced sensors and analytics, enabling intelligent systems that enhance energy efficiency and thermal comfort. This study introduces an innovative framework that combines multivariate analysis (MVA) with IoT technologies to improve indoor environmental wellness and optimize heating, ventilation, and air conditioning (HVAC) efficiency. It utilizes a nonlinear locally weighted regression (LWR) model, enhanced by sequential quadratic programming (SQP), to manage time-dependent variations in environmental factors such as temperature (Temp) and relative humidity (RH). The model was validated using 765 data points collected from three controlled indoor spaces in a university building. Its low computational overhead is compatible with real-time deployment on microcontroller-based platforms, making it well-suited for scalable and adaptive comfort control. Additionally, the framework contributes to more efficient HVAC operation by minimizing unnecessary energy consumption while maintaining occupant satisfaction.**

***Index Terms*—IoT, multivariate analysis, optimization, indoor wellness.**

I. Introduction

Efficient operation of heating, ventilation, and air conditioning (HVAC) systems is crucial for optimizing both energy consumption and occupants' comfort and well-being in modern buildings. In buildings, control schemes typically focus on single-variable feedback, regulating temperature (Temp) without simultaneously considering other environmental parameters that affect comfort and indoor air quality (IAQ)—such as relative humidity (RH), carbon dioxide (CO_2) concentration, and volatile organic compounds (VOCs) [1]. Recent studies address limitations by integrating internet of things (IoT) sensor networks with machine learning (ML)-based control algorithms. These systems employ diverse IoT sensors to monitor multiple environmental variables and use advanced methods, such as reinforcement learning, model predictive control, and supervised learning, to optimize HVAC operation for energy, comfort, and sometimes IAQ [2], [3].

However, the adoption of such advanced systems introduces significant computational challenges. Processing and analyzing high-frequency, high-dimensional IoT data can result in substantial computational load, especially for real-time optimization and control [4]. Complex algorithms, such as deep reinforcement learning or predictive models, may be difficult to deploy on resource-constrained edge devices or legacy building automation systems, often requiring cloud computing or powerful local hardware. These demands may limit scalability, increase latency, and create integration challenges in practical building contexts [5].

For the aforementioned reasons, this paper extends our earlier work [5] by using the same dataset but by enhancing the analytical framework and introducing completely new methodologies and results. While the previous study presented an IoT-based system combined with multivariate analysis (MVA), an assessment and prediction of indoor comfort conditions, this work refines that framework by integrating real-time control capabilities. It uses an optimization method to enhance indoor comfort by adjusting multiple environmental variables. It maps observations into principal component analysis (PCA) score space and applies sequential quadratic programming (SQP). Thus, by reducing the problem dimensionality with PCA, we enable the use of SQP efficiently in a lower-dimensional space to shift low-satisfaction points toward a high-satisfaction reference within physical constraints. This offers a novel alternative to traditional methods while keeping computational demands reasonable. Which allows multidimensional control beyond Temp, offering greater flexibility at the cost of increased memory and computation. Validated using Fanger's thermal comfort model [5], [6], the proposed algorithm significantly improves occupant satisfaction and reduces the predicted percentage of dissatisfied (PPD). By incorporating dynamic thermal modeling, occupant-specific adjustments, and advanced control strategies, this approach offers a computationally efficient and individualized solution for smart building wellness and HVAC optimization.

II. Framework Overview

A. IoT device

This study is predicated on extensive acquisitions performed with the Tauanito™ INAIR indoor air quality sensor from Taua. The sensor node is optimized for IoT-based environmental monitoring applications [7]. As outlined in the aforementioned study [5], the sensor node is capable of periodically acquiring Temp, pressure (P), RH, and IAQ parameters. The device also includes a GPS module for location tracking. Additionally, the integration of a GSM module facilitates reliable communications through cellular data transmission, a

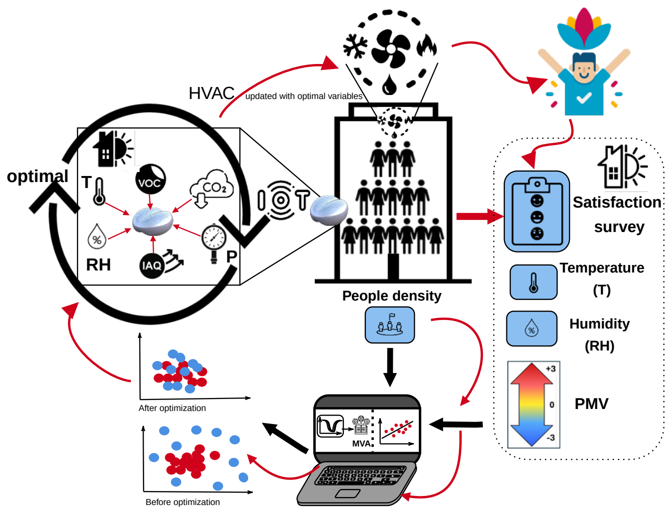


Fig. 1. Big picture of the proposed model

feature that is essential for seamless connectivity. An ultra-low power microcontroller unit (MCU) periodically acquires sensor data and transfers data to remote database servers through the Internet by exploiting a cellular data connection via global system for mobile communication (GSM). The remote servers include a MQTT server, a database server, and a web server to easily access data. To facilitate user accessibility, the ecosystem features a user-friendly web-based application, used in this work, which provides authenticated access to the acquired data via the Internet, ensuring General Data Protection Regulation (GDPR) compliance and incorporating an encryption layer.

B. Data collection and well-being assessment

The environmental data were retrieved from the IoT sensors distributed in three different rooms at the University of Bologna-Cesena Campus, including external factors such as Temp and RH gathered through the Visual Crossing database [8].

The present study is founded on the same dataset acquired and described in our previous work [5]. In summary, as described in detail in [5], three rooms (R_A, R_B, R_C), as shown in Fig 2, within a building of the University of Bologna were selected to accommodate two sensor nodes (S_1, S_2). S_1 was placed, alternatively, in R_A and R_B, where the occupation rate was approximately 100 in areas of 425 m^2 and 246 m^2, respectively, and were selected as objective rooms. S_2 was placed in R_C, where the peak occupation was approximately 26 individuals on 487 m^2, R_C was selected as the control room. A 21-day measurement campaign was conducted on working days, with data collected every 20 minutes and uploaded every two hours. Room occupancy was recorded four times per day. Participants received an automated questionnaire: 46 in room R_A, 27 in room R_B, and 16 in room R_C. All collected data were analyzed in an aggregated format.

Typically, two metrics are used to assess environmental well-being: the first metric is a self-rated overall satisfaction on a scale ranging from 1 (not satisfied) to 10 (perfectly satisfied); the second metric is the predicted mean vote (PMV), introduced by Fanger [6], which represents the average thermal sensation vote on a standard scale for a large group of individuals under various combinations of thermal environmental variables, activity, and clothing level [9]. More specifically, the PMV represents the average response of a significant number of individuals on a 7-point thermal sensation scale ranging from +3 (hot) to -3 (cold), with 0 indicating a neutral sensation, i.e. maximum wellness, [10], [11]. The model and equations for the calculation of PMV can be found in [12]–[14].

III. System model

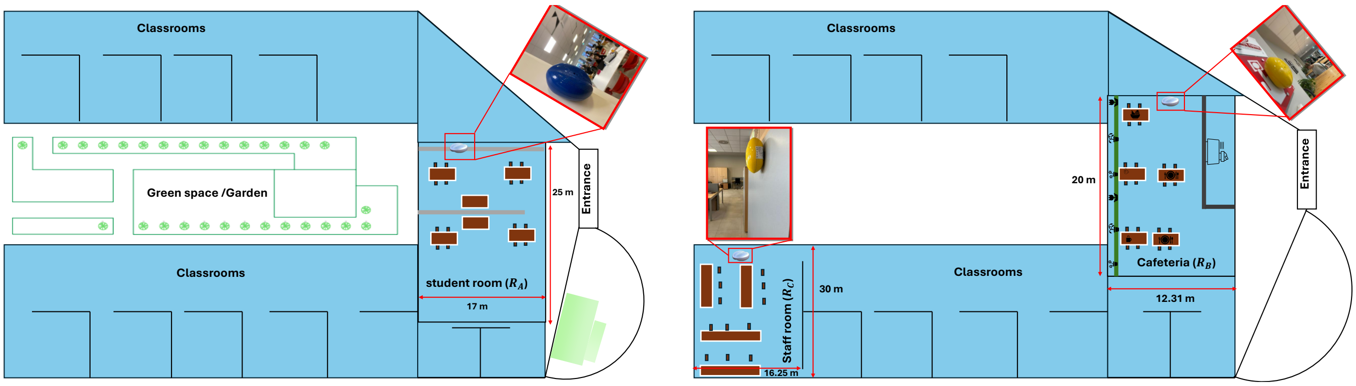
The proposed algorithm is based on the hypothesis that there is a region in the multidimensional score space where observations with high satisfaction are clustered. The analysis is conducted on a dataset comprising $N = 765$ observations.A reduced dataset, designated as $\boldsymbol{X}$, is constructed, encompassing the variables illustrated in Table I. Subsequent to this, two sub-datasets are derived from the aforementioned dataset. The first $\boldsymbol{X}_{ctr}$ contains all observations whose corresponding satisfaction levels are above a fixed (hyperparameter) threshold. A grid search was performed to optimize the satisfaction threshold for the algorithm, with results indicating ($= 7.5$) maximized prediction accuracy while maintaining computational feasibility The second $\boldsymbol{X}_{obj}$ contains the remaining observations. A locally weighted regression (LWR) [5], [15] model is built using 6 latent variables (root mean square error of calibration (RMSEC)$=$ 0.0641, root mean square error in prediction (RMSEP)$=$ 0.1564) calibrated onto the control dataset for the prediction of the satisfaction scores. The $\boldsymbol{X}_{obj}$ dataset is then applied to the model, and both the scores of the two datasets are extracted. The equation (1) shows how a generic single score point $\boldsymbol{T'}_{obj}$ $[1 \times A]$, linearly moves when the relative observation $\boldsymbol{X}_{obj}$ $[1 \times K]$ changes its own values by $\boldsymbol{\Delta}$ $[1 \times K]$. With K the number of variables and A is the number of te latent variables.

TABLE I
VARIABLE (VAR) IN $\boldsymbol{X}$

Var	Unit	Var	Unit
RH and External RH	[%]	CO_2 and VOC	[ppm]
External temperature (T)	$[°C]$	People density	$[\frac{\#\text{ people}}{\text{m}^2}]$
IAQ	[]	static IAQ	[]

$$
\begin{aligned}
\boldsymbol{T'}_{obj}(\boldsymbol{\Delta}) &= (\boldsymbol{X}_{obj} + \boldsymbol{\Delta}) \cdot L \\
&= \boldsymbol{X}_{obj} \cdot L + \boldsymbol{\Delta} \cdot L \\
&= \boldsymbol{T}_{obj} + \boldsymbol{\Delta} \cdot L
\end{aligned} \tag{1}
$$

With $\boldsymbol{T'}_{obj}$ representing the score of a single observation in the reduced-dimensional space, $\boldsymbol{L}$ is the loadings matrix that



(a) (b)

Fig. 2. On-site pictures of the INAIR devices on a simplified map of the Cesena campus: (a) Student Room (R_A) on the ground floor; (b) Cafeteria Room (R_B) and Staff Room (R_C) on the first floor.

maps the high-dimensional observation space to the lower-dimensional score space, and $\boldsymbol{T}_{obj}$ is the score vector corresponding to the observation $\boldsymbol{X}_{obj}$, and $\boldsymbol{\Delta}$ is the variation vector. The algorithm focuses on finding the variation vector $\boldsymbol{\Delta}$ that shifts all the objective scores towards the reference point $\boldsymbol{T}_{ref}$ $[1 \times A]$ computed as the mean value of the control ones. The shift vector $\boldsymbol{\Delta}$ is computed as the one for which the Euclidean distance between the reference and objective point is minimized or, equivalently, by computing the root of the orthogonality condition. The variation must respect the physical constraints; in particular, some variables, such as the external Temp and/or RH, are impossible to control directly; therefore, no changes should be allowed along these lines. The goal here is to find the change $\boldsymbol{\Delta}$ that moves the $\boldsymbol{T'}_{obj}$ as close as possible to the reference scores $\boldsymbol{T}_{ref}$. The vector $\boldsymbol{\Delta}$ is computed using a SQP approach. SQP is used for nonlinear optimization with constraints. Though computationally intensive, applying it in PCA-reduced space on an external server keeps the process efficient and scalable [16]. This approach involves linearizing the objective function and constraints and solving a quadratic subproblem at each step to find the optimal $\boldsymbol{\Delta}$. Once computed, $\boldsymbol{\Delta}$ is added to the original objective observation $\boldsymbol{X}_{\text{obj}}$, thereby modifying all the environmental variables' levels while ensuring that they stay within the specified physical constraints. For each observation in the data $\boldsymbol{X}_{\text{obj}}$ is adjusted using the LWR model.

The model is generated using the LWR model on a control dataset $\boldsymbol{X}_{\text{ctr}}$, which produces the model's scores and loadings. The reference score $\boldsymbol{T}_{\text{ref}}$ is computed as the mean of the scores from the model. The loadings matrix $\boldsymbol{L}$ is extracted from the model, which defines the relationship between the observed data and the principal components. After optimization, the adjusted observation $\boldsymbol{X}_{\boldsymbol{new}}$ is computed by adding the change vector to the original observation.

IV. Discussion

To check the effectiveness of the methods, a dataset $\boldsymbol{X}^{cc}$ is built using the variables (External and indoor Temp and RH, VOC, CO_2, IAQ, static IAQ) from which a training $\boldsymbol{X}^{cc}_{cal}$ and test $\boldsymbol{X}^{cc}_{test}$ set are obtained with a $80\% - 20\%$ ratio with respect to the original data. A partial least square (PLS) model is calibrated onto the calibration set and validated on the test set for the prediction of PMV scores. Once the model is built, a new dataset $\boldsymbol{X}^{opt}$ is built in the same fashion as $\boldsymbol{X}^{cc}$, but the Temp variable is replaced with the optimal one, which obtained through the proposed algorithm, and the modified dataset is applied to predict the optimal PMV. Ideally, the newly predicted scores are all equal, or close, to zero. Each new predicted PMV whose value falls in the ± 0.5 range contributes to the estimation of the overall satisfaction percentage evaluated as the area under Fanger's curve between the set values.

The predicted (optimal) Temp versus the actual indoor Temp is shown in Fig 3. The actual indoor Temp, which ranges between 19°C and 29°C, with the predicted optimal Temp range of 19°C to 26°C, is designed to maximize occupant satisfaction. The actual Temp fluctuates outside the ideal comfort zone, especially in environments where the Temp exceeds 25°C or falls below 19°.

Fig 4 illustrates the observations both before and after the algorithm was applied, with the optimal predicted Temp, where the target points are close to the reference. The red points are above the threshold and are used to calculate the reference point; the blue points are the target points to be improved. It demonstrates how the algorithm effectively moves unsatisfactory conditions closer to the optimal comfort zone, validating its capability to improve environmental wellness through multidimensional adjustments.

The effect of the algorithm on thermal comfort is illustrated in Fig. 5, which compares the distributions of PMV scores and their corresponding PPD before and after algorithm application. The original distribution displays a wider spread, with a significant proportion of PMV values falling outside the optimal comfort range. After applying the algorithm, the distribution becomes more concentrated around thermal neutrality, with a mean PMV of -0.27 and a standard deviation of 0.22.

The predicted satisfaction, defined as the proportion of occupants within the PMV comfort range of ± 0.5, increases significantly, reaching 85.51%. Correspondingly, the PPD remains below 10%, which meets the recommended threshold defined in ASHRAE Standard 55 and ISO 7730 [6], [10], [11]. These results clearly demonstrate the algorithm's effectiveness in enhancing indoor thermal comfort and aligning environmental conditions with human thermal preference.

In fact, keeping indoors Temp between 19°C and 24°C aligns with recognized thermal comfort guidelines and supports energy efficiency. As stated in ISO 7730:2023 [11], this Temp range is ideal for occupant comfort. Furthermore, the European Union's "Playing my part" campaign highlights that lowering indoor heating by just 1°C can lead to a 7–15% reduction in HVAC energy consumption [17]. Following these standards allows individuals to maintain comfort while actively supporting energy-saving initiatives.

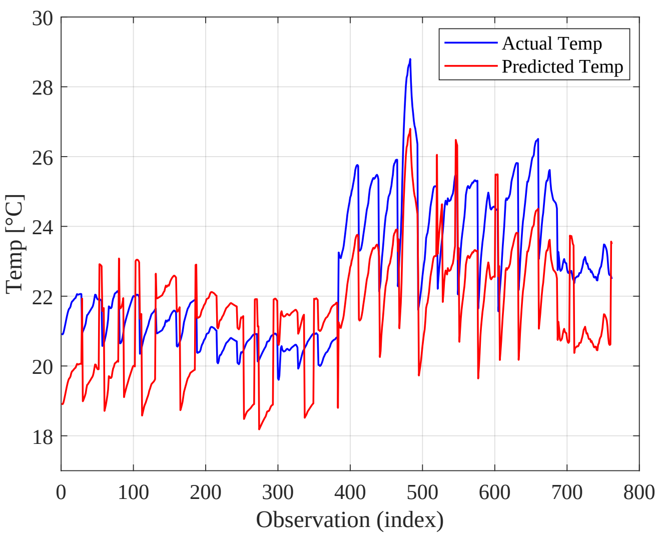


Fig. 3. Prediction of optimal Temp with respect to the original measured one.

V. Conclusion

The objective of this paper is to enhance environmental wellness through an integrated approach combining IoT technologies with MVA. For the first time in this context, we employ a pipeline composed of variable reduction, a LWR model, and constrained nonlinear optimization using the high-precision SQP algorithm, demonstrating its effectiveness in handling complex environmental control scenarios. Using MVA techniques, including the proposed algorithm and LWR model, allows for precise predictions of optimal Temp and other variables settings that enhance occupant satisfaction while ensuring energy-efficient HVAC operations. The observed reduction in PPD values validates the model's ability to maintain thermal comfort within established ISO standards. The framework's adaptability makes it a promising solution for IoT-driven intelligent building management systems.

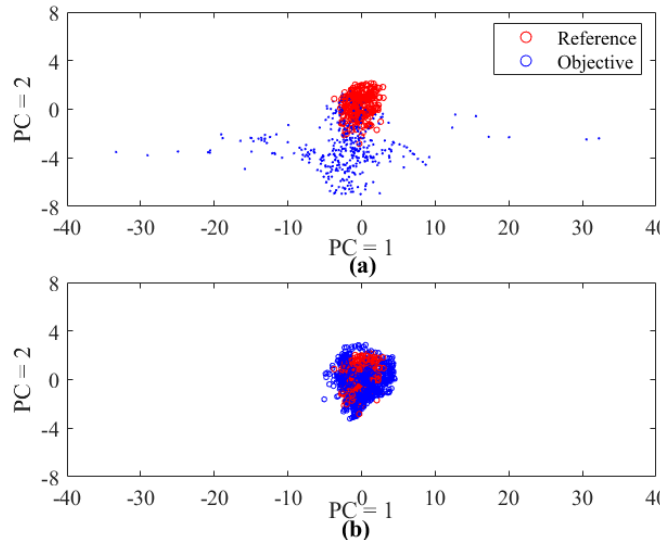


Fig. 4. Score plots of points (a) before and (b) after the algorithm application.

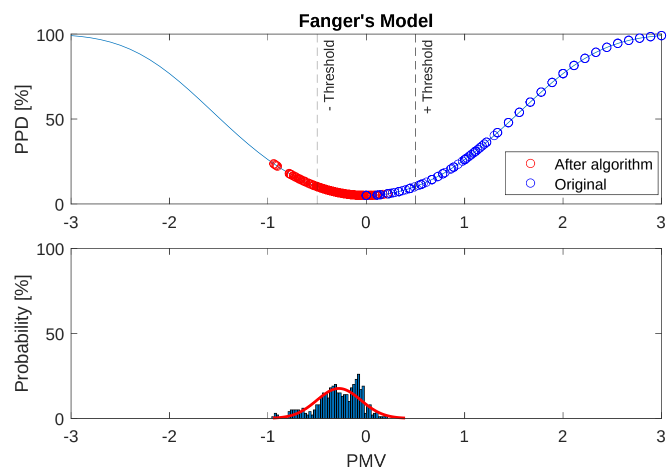


Fig. 5. Predicted PPD scores using the Fanger's model onto the predicted PMV.

Acknowledgment

This study was supported by the Agritech National Research Center and received funding from the European Union NextGenerationEU (PNRR) – Mission 4 Component 2, Investment 1.4 – D.D. 1032 17/06/2022, CN00000022. The authors would like to acknowledge the financial support by the Chips Joint Undertaking (Chips JU) and its members under grant agreement No. 101112286 (LoLiPoP IoT).

The authors would like to express their special thanks to Marco Bellomo and Andrea Romagnoli from Taua Società Benefit for their valuable support and insightful contributions to this work.

References

[1] S. Xia, P. Wei, Y. Liu, A. Sonta, and X. Jiang, "A multi-task deep reinforcement learning-based recommender system for co-optimizing energy, comfort, and air quality in commercial buildings with humans-in-the-loop," *Data-Centric Engineering*, vol. 5, p. e26, Nov. 2024.

[2] A. Morteza, H. K. Nazari, and P. Pahlevani, "An iot framework for building energy optimization using machine learning-based mpc," 2024. [Online]. Available: https://arxiv.org/abs/2408.13294

[3] J. Lee, D.-O. Woo, J. Jang, L. Junghans, and S.-B. Leigh, "Collection and utilization of indoor environmental quality information using affordable image sensing technology," *Energies*, vol. 15, no. 3, p. 921, Jan. 2022.

[4] O. Afif *et al.*, "A versatile, machine-learning-enhanced rf spectral sensor for developing a trunk hydration monitoring system in smart agriculture," *Sensors*, vol. 24, no. 19, p. 6199, Sept. 2024.

[5] O. Afif, G. Ingenito, M. Bellomo, A. Romagnoli, M. Tartagni, and A. Romani, "Iot-based indoor thermal comfort prediction using multivariate statistical analysis," *IEEE Sensors Journal*, vol. 25, no. 8, pp. 14 358–14 369, Mar. 2025.

[6] P. O. Fanger *et al.*, "Thermal comfort. analysis and applications in environmental engineering." *Thermal comfort. Analysis and applications in environmental engineering.*, 1970.

[7] Tāua s.r.l. società benefit, "Tāua - internet of things and web of things solutions," 2024. [Online]. Available: https://www.taua.it/en/

[8] V. C. Corporation, "Visual crossing weather data services," 2024. [Online]. Available: https://www.visualcrossing.com/

[9] J. Van Hoof, M. Mazej, and J. L. Hensen, "Thermal comfort: research and practice," *Frontiers in Bioscience-Landmark*, vol. 15, no. 2, pp. 765–788, 2010.

[10] ASHRAE, "Thermal environmental conditions for human occupancy," 2023. [Online]. Available: https://www.ashrae.org/technical-resources/bookstore/standard-55-thermal-environmental-conditions-for-human-occupancy.

[11] International Organization for Standardization, "Ergonomics of the thermal environment - analytical determination and interpretation of thermal comfort using calculation of the pmv and ppd indices and local thermal comfort criteria," https://webstore.ansi.org/standards/iso/isodis77302023, 2023, iSO/DIS 7730:2023.

[12] S. Zare, N. Hasheminezhad, K. Sarebanzadeh, F. Zolala, R. Hemmatjo, and D. Hassanvand, "Assessing thermal comfort in tourist attractions through objective and subjective procedures based on iso 7730 standard: A field study," *Urban climate*, vol. 26, pp. 1–9, Dec. 2018.

[13] B. E. Ainsworth, W. L. Haskell, A. S. Leon, D. R. Jacobs Jr, H. J. Montoye, J. F. Sallis, and R. S. Paffenbarger Jr, "Compendium of physical activities: classification of energy costs of human physical activities." *Medicine and science in sports and exercise*, vol. 25, no. 1, pp. 71–80, 1993.

[14] B. E. Ainsworth *et al.*, "Compendium of physical activities: an update of activity codes and met intensities," *Medicine and science in sports and exercise*, vol. 32, no. 9; SUPP/1, pp. S498–S504, 2000.

[15] W. S. Cleveland and S. J. Devlin, "Locally weighted regression: an approach to regression analysis by local fitting," *Journal of the American statistical association*, vol. 83, no. 403, pp. 596–610, 1988.

[16] P. T. Boggs and J. W. Tolle, "Sequential quadratic programming," *Acta numerica*, vol. 4, pp. 1–51, 1995.

[17] European Commission, "Playing my part - energy prices and consumer actions," 2023, accessed: 2025-03-20. [Online]. Available: https://energy.ec.europa.eu/topics/markets-and-consumers/actions-and-measures-energy-prices/playing-my-part-en