

Journal Scientific and Applied Research, Vol. 30, 2026
International Journal

ISSN 1314-6289 (Print)
ISSN 2815-4622 (Online)

Original Contribution

INTELLIGENT HEAT SOURCE SELECTION UNDER STREAMING DATA QUALITY AND OPERATIONAL RISK CONSTRAINTS

Milen St. Davidkov

Department of Communication and Computer Engineering and Security Technologies, Faculty of Technical Sciences, Konstantin Preslavsky University of Shumen, Shumen 9712, 115 Universitetska St., e-mail: milen_davidkov@abv.bg

Abstract: *This paper proposes a quality-aware decision model for selecting heat sources in an integrated thermal-energy system. The method combines technical, economic and environmental utility with the quality of streaming measurements and an FMEA-based operational-risk penalty. A source is selected only after critical data-quality constraints are satisfied; otherwise, the algorithm requests redundant confirmation or activates a validated safe mode. A reproducible computational experiment with five heat-source alternatives and five disturbance scenarios compares the proposed method with a quality-blind baseline. Under normal conditions both methods select the same source. Under delayed, drifting, lost and combined data disturbances, the proposed model rejects the unreliable nominal optimum and selects an alternative with verifiable input data. The model therefore links data-quality assessment directly to an explainable operational decision rather than treating it as a separate monitoring task.*

Key Words: *Integrated thermal-energy system, Streaming data quality, Heat-source selection, Operational risk, Multicriteria decision making.*

1. Introduction

Integrated thermal-energy systems combine dispatchable and intermittent sources, storage and demand-side flexibility. Heat pumps, biomass and gas boilers, electric boilers, solar thermal collectors, geothermal sources, waste heat and thermal storage may participate in a common supply system. Their coordinated use can reduce cost and emissions, but the operating decision depends on measurements of demand, temperature, flow, pressure, source availability and equipment condition. ISO 50001 places monitoring and measurement within the energy-management process [1], while smart-energy research emphasizes the coordinated use of multiple sectors and resources [2].

Conventional dispatch usually assumes that every input value is sufficiently reliable. This assumption is unsafe when measurements arrive late, packets are lost, units are inconsistent or a sensor drifts. Data quality is multidimensional and context dependent [3]–[5]. In an industrial control system, low-quality information can alter source ranking even when the optimization procedure itself is mathematically correct. The relevant problem is

therefore not only which source has the highest nominal utility, but whether the information supporting that ranking is trustworthy enough for automatic action.

The objective of this study is to integrate streaming data quality and operational risk directly into heat-source selection. The contribution is a constrained multicriteria score in which unreliable information reduces the utility of an alternative and increases its risk penalty. A four-level decision policy separates unrestricted optimization, restricted operation, redundant confirmation and safe fallback. The model is evaluated in a controlled scenario-based computational experiment whose normalized inputs are reported in full.

2. Related work

Multicriteria decision-making methods are widely applied in sustainable-energy planning because technical, economic, environmental and social objectives cannot be reduced to a single physical measure [6], [7]. These methods support transparent ranking, but they normally treat criterion values as known. In streaming industrial environments, the confidence in those values changes with communication delay, missing observations and equipment condition. The distinction between intrinsic data properties and fitness for use is central to established data-quality models [3–5].

Previous work by Denev and Tsankov discusses data quality and processing in smart devices [8]. Their studies of visual PLC monitoring and field-level industrial networks show that packet organization, communication delay and the structure of the industrial network affect the availability of current process information [9], [10]. These findings support the placement of structural and time checks close to the PLC or edge device instead of transferring every raw value to a remote layer.

Risk is represented through failure mode and effects analysis in accordance with the general procedure standardized in IEC 60812 [11]. The earlier FMEA application by Arabadzhieva-Kalcheva and Tsankov provides a basis for linking failure causes to severity, occurrence and detectability [12]. Cause-consequence analysis adds the connection between an initiating event, protective action and final system state [13]. Nonlinear transformations are useful when a gradual deterioration of an input must produce a progressively stronger penalty [14]. Stability-oriented diagnostic reasoning also supports the use of a conservative fallback when control information is uncertain [15].

3. Proposed method

3.1. Data-quality vector

For heat source s at time t, the required input is represented by a quality vector containing accuracy A, completeness C, timeliness T, validity V, consistency K, provenance P and reliability R. Each component is normalized to the interval from zero to one:

$$Q_s(t) = [A_s(t), C_s(t), T_s(t), V_s(t), K_s(t), P_s(t), R_s(t)] \in [0,1]^7 \quad (1)$$

The aggregated quality $q_s(t)$ is calculated as a weighted sum of the seven components of the quality vector. The coefficients v_j express their relative importance and their sum is equal to one. Depending on the operating conditions, greater weight may be assigned to timeliness and reliability or to accuracy and consistency.

$$q_s(t) = \sum_{j=1}^{7} v_j q_{j,s}(t), \quad \sum_{j=1}^{7} v_j = 1 \quad (2)$$

Here, $q_{j,s}(t)$ denotes the j-th component of the data-quality vector for source s, while v_j is its corresponding weight.

3.2. Nominal source utility

The nominal utility U_s combines normalized economic, environmental, technical and availability criteria. Higher values are preferable. Cost and emissions are converted to benefit form by subtracting their normalized values from one. The criterion weights sum to one and may be selected by the operator or obtained through an established multicriteria procedure.

$$U_s = a_1(1 - c_s) + a_2(1 - e_s) + a_3\eta_s + a_4 g_s, \quad \sum_{k=1}^{4} a_k = 1 \tag{3}$$

Here c_s is normalized operating cost, e_s is normalized emissions intensity, η_s is conversion efficiency or coefficient-of-performance utility, and g_s is technical readiness and available capacity. The score is evaluated only for sources that can satisfy the required thermal output or participate in a feasible combination. The heat balance remains a hard constraint:

$$\sum_s H_s \geq H_{(dem)}, \quad 0 \leq H_s \leq y_s H_{(s,max)}, \quad y_s \in \{0,1\} \tag{4}$$

Here, H_s is the thermal output of source s, H_{dem} is the required thermal demand, $H_{s,max}$ is the maximum available output, and y_s is a binary variable indicating whether the source is activated.

3.3. Operational risk and quality-aware score

$$r_s = Sev_s \cdot Occ_s \cdot NDet_s \tag{5}$$

For each source, the normalized FMEA risk index is defined in terms of severity Sev_s, occurrence Occs and non-detectability $NDet_s$. The multiplicative form preserves the logic of the risk-priority number while producing a value between zero and one.

The proposed quality-aware score J_s reduces the nominal utility according to the quality of the available data and introduces a nonlinear penalty that combines operational risk with missing confidence. The coefficient $\beta \geq 0$ determines the influence of the risk penalty, while the exponent $\gamma \geq 1$ makes this penalty increase more rapidly as data quality deteriorates

$$J_s = q_s U_s - \beta r_s (1 - q_s)^{\gamma} \tag{6}$$

Let Ω denote the set of heat sources that satisfy the thermal-demand constraint, the operating limits and the minimum data-quality requirements. The selected source s* is the feasible alternative with the highest quality-aware score, as illustrated in the decision process (Fig. 1):

$$s^* = \arg\max_{(s \in \Omega)} J_s \tag{7}$$

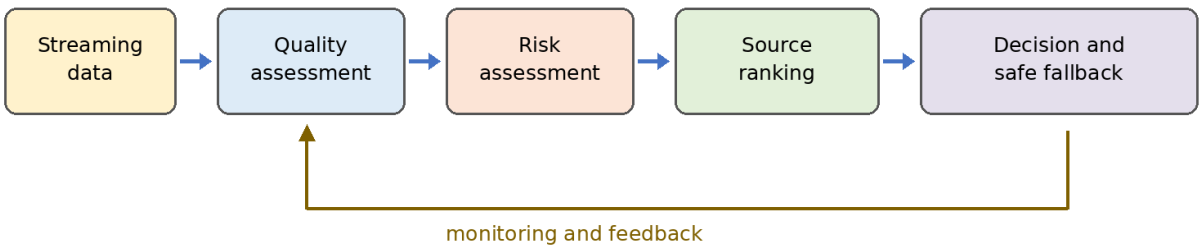


Fig. 1. Quality-aware heat-source selection and feedback loop

For the present computational experiment, the quality thresholds were defined by the author as illustrative operational limits (Table 1). In a practical implementation, these thresholds should be calibrated according to sensor specifications, communication requirements, historical data and the safety requirements of the particular thermal-energy system.

4. Computational experiment

A deterministic scenario-based experiment was implemented to verify the decision logic. It is not presented as a field measurement. Five alternatives were considered: solar thermal with storage (ST), heat pump (HP), biomass boiler (BB), gas boiler (GB) and electric boiler (EB). The baseline selects the largest nominal utility U_s and ignores quality. The proposed method uses equation (6) with $\beta=0.25$ and $\gamma=1$. For the controlled computational experiment, the author assigned the normalized illustrative risk indices [0.25, 0.30, 0.40, 0.55, 0.20] to [ST, HP, BB, GB, EB], respectively. These values are used to demonstrate the decision logic and are not derived from a field-based FMEA assessment.

Five scenarios represent normal operation, delayed solar data, heat-pump sensor drift, loss of gas-boiler communication and a combined disturbance. Table 2 gives the complete vectors used in the calculations. Values are ordered [ST, HP, BB, GB, EB]. They are normalized hypothetical scenario parameters defined by the author to test the decision rules under controlled disturbances. They are not field measurements, measured efficiencies or market prices.

Table 1. Data-quality classes and decision policy

Quality class	Condition	Permitted decision	System response
I - high	$q_s \geq 0.85$ and all critical limits met	unrestricted ranking	automatic dispatch
II - acceptable	$0.70 \leq q_s < 0.85$	restricted ranking	increased monitoring and reserve
III - low	$0.50 \leq q_s < 0.70$	non-critical action only	redundant confirmation or correction
IV - critical	$q_s < 0.50$ or critical limit violated	no optimization	validated safe mode and alarm

Table 2. Reproducible normalized inputs for the five scenarios

Scenario	Nominal utility vector U	Quality vector q	Disturbance
Normal	[0.86, 0.82, 0.73, 0.69, 0.58]	[0.96, 0.95, 0.93, 0.96, 0.95]	none
Delayed solar	[0.86, 0.82, 0.73, 0.69, 0.58]	[0.45, 0.95, 0.93, 0.96, 0.95]	late and incomplete ST data
HP sensor drift	[0.70, 0.90, 0.78, 0.77, 0.65]	[0.92, 0.48, 0.91, 0.94, 0.95]	biased HP temperature signal
Gas communication loss	[0.68, 0.82, 0.76, 0.91, 0.70]	[0.90, 0.94, 0.92, 0.35, 0.96]	missing GB packets
Combined disturbance	[0.88, 0.86, 0.78, 0.80, 0.69]	[0.55, 0.60, 0.88, 0.52, 0.94]	multiple degraded streams

All calculations were performed directly from the reported vectors. For example, in the delayed-solar scenario the solar score is 0.45×0.86-0.25×0.25×(1-0.45)=0.353, while the heat-pump score is 0.95×0.82-0.25×0.30×(1-0.95)=0.775. The baseline therefore retains ST, whereas the quality-aware method selects HP.

5. Results and discussion

Under normal conditions both methods select solar thermal, which shows that the quality term does not change a well-supported nominal optimum. In each disturbed scenario, however, the baseline continues to select the alternative with the largest U value although the supporting stream is unreliable. The proposed method changes the decision in all four disturbance cases (Table 3). This is the intended behavior: the method does not claim that HP, GB or BB is universally superior; it selects the best supported feasible alternative for the current state.

The largest score reduction occurs under gas communication loss. The nominally attractive gas boiler obtains J=0.229 because q=0.35 and r=0.55. The heat pump reaches 0.766 and becomes the preferred source. In the combined scenario, biomass is selected although the electric boiler has the highest quality. This result follows from the balance between evidence quality and nominal technical-economic utility; data quality is influential but does not replace the energy criteria.

Table 3. Comparison of baseline and quality-aware decisions

Scenario	Baseline selection	Quality-aware selection	Winning score J	Interpretation
Normal	ST	ST	0.823	same decision under reliable data
Delayed solar	ST	HP	0.775	unreliable nominal optimum rejected
HP sensor drift	HP	GB	0.716	biased HP input excluded
Gas communication loss	GB	HP	0.766	missing GB data triggers alternative
Combined disturbance	ST	BB	0.674	robust source with better evidence selected

In the delayed-solar, HP-sensor-drift and gas-communication-loss scenarios, the nominally preferred source violates the critical data-quality requirement and is therefore excluded. In the combined-disturbance scenario, solar thermal remains admissible but obtains a lower quality-aware score than biomass. The hard quality threshold excludes critically unreliable alternatives, while the continuous quality-risk term differentiates among the admissible sources.

The method is compatible with a layered implementation. Structural validation and timestamps can be checked at PLC or edge level, cross-source consistency and short-window statistics can be computed in a fog node, and long-term models can be maintained centrally.

Such distribution follows established edge-computing principles [16], [17] and the field-level architecture in [10]. Packet optimization [9] is relevant when the monitoring rate is adapted to the current quality class.

The experiment has limitations. The normalized values are controlled inputs rather than measurements from an operating district-heating system. The illustrative risk indices are not based on a complete field FMEA and should be calibrated through expert assessment and operational failure data before practical implementation. Source start-up time, minimum load, ramp rate and multi-period storage dynamics are represented only through U and feasibility. The next validation stage should use measured time series and solve a multi-period dispatch problem. Nevertheless, the complete scenario vectors make the present decision results reproducible and demonstrate the effect of the proposed quality-risk coupling without fabricating field evidence.

6. Conclusion

A quality-aware model for intelligent heat-source selection was developed. It combines nominal multicriteria utility, a seven-component streaming-data quality assessment, an FMEA risk index and critical admissibility thresholds. A controlled experiment with five alternatives and five scenarios showed identical baseline and proposed decisions under reliable data and different decisions under four disturbances. In every disturbed case the method rejected an unreliable nominal optimum and selected a feasible alternative supported by better-quality information.

The contribution is the direct integration of data confidence and operational risk into source ranking and safe fallback. The method is transparent, computationally simple and suitable for PLC-edge-fog implementation. Future work will calibrate the parameters with measured thermal-system data, extend the heat-balance model over multiple periods and compare the decisions with established robust and stochastic optimization methods.

Acknowledgments

The author thanks the Faculty of Technical Sciences at Konstantin Preslavsky University of Shumen for the academic environment supporting this study.

References:

[1] ISO 50001:2018, Energy management systems - Requirements with guidance for use. International Organization for Standardization, Geneva, 2018.

[2] Lund, H., Østergaard, P. A., Connolly, D., Mathiesen, B. V., Smart energy and smart energy systems. Energy, vol. 137, 2017, pp. 556-565. DOI: 10.1016/j.energy.2017.05.123.

[3] ISO/IEC 25012:2008, Software engineering - Software product Quality Requirements and Evaluation (SQuaRE) - Data quality model. ISO, Geneva, 2008.

[4] Wang, R. Y., Strong, D. M., Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems, vol. 12, no. 4, 1996, pp. 5-33. DOI: 10.1080/07421222.1996.11518099.

[5] Batini, C., Cappiello, C., Francalanci, C., Maurino, A., Methodologies for data quality assessment and improvement. ACM Computing Surveys, vol. 41, no. 3, 2009, Article 16. DOI: 10.1145/1541880.1541883.

[6] Pohekar, S. D., Ramachandran, M., Application of multi-criteria decision making to sustainable energy planning - A review. Renewable and Sustainable Energy Reviews, vol. 8, no. 4, 2004, pp. 365-381. DOI: 10.1016/S1364-0321(03)00070-1.
[7] Mardani, A., Jusoh, A., Zavadskas, E. K., Cavallaro, F., Khalifah, Z., Sustainable and renewable energy: An overview of the application of multiple criteria decision making techniques and approaches. Sustainability, vol. 7, no. 10, 2015, pp. 13947-13984. DOI: 10.3390/su71013947.
[8] Denev, D. R., Tsankov, Ts. S., A survey of data quality and processing of information in innovative smart devices. 2022 International Conference Automatics and Informatics (ICAI), IEEE, 2022, pp. 311-315, ISBN 978-1-6654-7625-6.
[9] Denev, D. R., Tsankov, Ts. S., Communication-based dynamic packet optimization for visual PLC monitoring. Annual University Scientific Conference of the National Military University Vasil Levski, Veliko Tarnovo, 2022, ISSN 2367-7481, pp. 1259-1266.
[10] Denev, D. R., Tsankov, Ts. S., Design and implementation of a field level of an industrial communication network for the needs of the industry. ICEST 2022, Ohrid, 2022, pp. 1-4. DOI: 10.1109/ICEST55168.2022.9828607.
[11] IEC 60812:2018, Failure modes and effects analysis (FMEA and FMECA). International Electrotechnical Commission, Geneva, 2018.
[12] Arabadzhieva-Kalcheva, N. A., Tsankov, Ts. S., Failure modes and effects analysis - FMEA. Proceedings of the Scientific Conference MATTECH 2022, vol. 2, 2022, pp. 58-62.
[13] Arabadzhieva-Kalcheva, N. A., Tsankov, Ts. S., Todorova, M. P., Cause-consequence analysis as a risk management method. Annual of the Faculty of Technical Sciences, Shumen University, vol. XIV E, 2024, pp. 184-188. DOI: 10.46687/YT24NEL.
[14] Tsankov, Ts. S., Applications of nonlinear functions in communication systems. Shumen, 2018, 204 p., ISBN 978-619-201-212-0.
[15] Ahmedov, I. S., Chirakov, D. P., Tsankov, Ts. S., Inspection of automotive controllers for stability via communication systems with specialized diagnostic software. Journal of Physics and Technology, 2018, pp. 52-55.
[16] Iorga, M., Feldman, L., Barton, R., Martin, M. J., Goren, N. S., Mahmoudi, C., Fog Computing Conceptual Model. NIST Special Publication 500-325, 2018. DOI: 10.6028/NIST.SP.500-325.
[17] Shi, W., Cao, J., Zhang, Q., Li, Y., Xu, L., Edge computing: Vision and challenges. IEEE Internet of Things Journal, vol. 3, no. 5, 2016, pp. 637-646. DOI: 10.1109/JIOT.2016.2579198.