

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

ÉCOLE NATIONALE POLYTECHNIQUE



Département d'Électronique

Entreprise d'accueil : Holcim El-Djazaïr – Oggaz

End-of-Study Project Dissertation submitted for the
State Engineer's Degree in Electronics

Design of an Embedded Soft Sensor for Estimating the Apparent Calorific Value of Alternative Solid Fuels in a Cement Plant

Presented by :

Oussama ALLAOUA & Mohamed El Amine BOUGDAH

Presented and publicly defended on 06/07/2026

Jury Members :

President	Dr. Messaoud CHAKIR	ENP
Supervisor	Dr. Nora BOUMELA	ENP
Examiner	Dr. Fatiha LANI	ENP

ENP 2026

ENP, 10 Rue des Frères OUDEK, El Harrach 16200, Alger, Algérie

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

ÉCOLE NATIONALE POLYTECHNIQUE



Département d'Électronique

Entreprise d'accueil : Holcim El-Djazaïr – Oggaz

End-of-Study Project Dissertation submitted for the
State Engineer's Degree in Electronics

Design of an Embedded Soft Sensor for Estimating the Apparent Calorific Value of Alternative Solid Fuels in a Cement Plant

Presented by :

Oussama ALLAOUA & Mohamed El Amine BOUGDAH

Presented and publicly defended on 06/07/2026

Jury Members :

President	Dr. Messaoud CHAKIR	ENP
Supervisor	Dr. Nora BOUMELA	ENP
Examiner	Dr. Fatiha LANI	ENP

ENP 2026

ENP, 10 Rue des Frères OUDEK, El Harrach 16200, Alger, Algérie

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

ÉCOLE NATIONALE POLYTECHNIQUE



Département d'Électronique

Entreprise d'accueil : Holcim El-Djazaïr – Oggaz

Mémoire de Projet de Fin d'Études pour l'obtention du
Diplôme d'Ingénieur d'État en Électronique

Conception d'un capteur logiciel embarqué pour l'estimation du pouvoir calorifique apparent des CSA en cimenterie

Présenté par :

Oussama ALLAOUA & Mohamed El Amine BOUGDAH

Présenté et soutenu publiquement le 06/07/2026

Composition du Jury :

Président	Dr. Messaoud CHAKIR	ENP
Encadrante	Dr Nora BOUMELA	ENP
Examineur	Dr Fatiha LANI	ENP

ENP 2026

ENP, 10 Rue des Frères OUDEK, El Harrach 16200, Alger, Algérie

تبحث هذه الأطروحة في تقدير القيمة الحرارية الظاهرية للوقود الصلب البديل في مكبس مصنع إسمنت، بالاعتماد فقط على إشارات التشغيل المتاحة في نظام التحكم. ولأن هذه القيمة لا تُقاس أنياً، بل تُعرف لاحقاً من خلال تحليل مخبري لعينة يومية، تبدأ الدراسة بتقييم قابلية رصدها، ثم تطوّر تواتراً رقمياً معياراً لتوليد أحداث انتقالية واقعية وموسومة. وتُظهر النتائج أن إشارة قابلة للاسترجاع موجودة بالفعل، لكنها تظل ضعيفة ومحجوبة جزئياً بفعل التحكم في حلقة مغلقة؛ كما أن أفضل أداء يتحقق باستخدام نموذج خطي مدعوم بهندسة ميزات مناسبة، في حين يظل مسار ميزان الطاقة شديد الحساسية لعدم اليقين في قياس الكتلة.

الكلمات المفتاحية: القيمة الحرارية الظاهرية؛ الوقود الصلب البديل؛ صناعة الإسمنت؛ التوامم الرقمي؛ الرصدية؛ المستشعر البرمجي.

Résumé

Cette thèse étudie l'estimation de la valeur calorifique apparente d'un combustible solide de substitution dans le calcinateur d'une cimenterie, à partir des seuls signaux déjà disponibles dans le système de conduite. Comme cette valeur n'est accessible qu'a posteriori à partir d'une analyse quotidienne en laboratoire, l'étude évalue d'abord son observabilité, puis développe un jumeau numérique calibré afin de générer des événements transitoires réalistes et étiquetés. Les résultats montrent qu'un signal récupérable existe, mais qu'il demeure faible et partiellement masqué par la régulation en boucle fermée ; un modèle linéaire associé à une ingénierie de variables adaptée offre les meilleures performances, tandis que l'approche fondée sur le bilan énergétique reste fortement sensible à l'incertitude sur la mesure de la masse.

Mots-clés : valeur calorifique apparente ; combustible solide de substitution ; cimenterie ; jumeau numérique ; observabilité ; capteur virtuel.

Abstract

This thesis investigates the estimation of the apparent calorific value of alternative solid fuel in a cement calciner using only process signals available in the plant control system. Because this quantity is only available after daily laboratory analysis, the work first evaluates its observability, then develops a calibrated digital twin to generate realistic, labeled transient events. The results show that a recoverable signal exists, although it remains weak and partially masked by closed-loop control. A linear model with appropriate feature engineering achieves the best performance, while the energy-balance approach remains highly sensitive to uncertainty in fuel mass measurement.

Keywords: apparent calorific value; alternative solid fuel; cement plant; digital twin; observability; soft sensor.

Acknowledgements

We would like to express our sincere gratitude to **Dr Boumela**, whose guidance, encouragement, and high standards have shaped this work from its earliest stages to its completion. Beyond her academic supervision, we are especially thankful for the effort she invested in making this internship opportunity possible and for the confidence she placed in us throughout the project. Her availability, constructive feedback, and insistence on scientific rigour have been invaluable, and they have contributed greatly to both this thesis and our personal development.

We would also like to thank **Holcim El-Djazaïr** and the entire team at the **Oggaz cement plant** for welcoming us into their facility and giving us the opportunity to work on a real industrial problem. Our thanks go in particular to:

- **Mr. Abdelkader AZZOUZ**, the OT (Operational Technology) Responsible and our supervisor at the plant, whose support was central to this project. We are deeply grateful for his patience in explaining the DCS architecture and the control loops, for his constant availability, and for the guidance he provided at every stage of our work.
- **Mr. Mohammed Nadir KHELOUFI**, Automation Level 2 Engineer, for retrieving and extracting the historical process data from the TIS historian system, and for taking a considerable amount of work off our shoulders throughout the internship.
- **Mr. Mahmoud HARICHEN**, Geocycle Engineer, for his insight into the laboratory measurements and for helping us understand the chemistry of the alternative fuels and their effects on the process.
- **Ms. Salima DJADEL** and **Ms. Silia BENKACI**, from the Geocycle laboratory, for providing the laboratory data and for their valuable insight on how to clean it.
- **Mr. Benabdelkader ARABI**, R&D Director, for sharing his strategic vision on decarbonization and alternative fuels.
- **Ms. Nawel KOTNI**, Training Manager at the Oggaz plant, for her administrative support and for coordinating our internship logistics.
- The production engineers and control-room operators who generously shared their time, technical expertise, and daily operational insights with us.

Finally, Oussama would like to express his deep gratitude to the **Polytech** community for welcoming him back with kindness, for including him throughout this journey, and for giving him the opportunity to complete his degree. This support meant a great deal to him, and he will always remember it with appreciation.

Dedication

To my parents, whose love, sacrifices, and unwavering belief in me have made every achievement possible.

To my brothers and sisters, for their constant encouragement, support, and the strength that comes from knowing I could always count on them.

To my family, my relatives, and my friends, who have stood by me throughout this journey and shared both its challenges and its successes.

This work is dedicated to all of you, with my deepest gratitude.

Oussama

To my parents, for their endless patience, love, and support throughout this journey.

To my family and loved ones, for their encouragement, trust, and presence at every step.

This work is dedicated to you with sincere appreciation and gratitude.

Mohamed

Table of Contents

List of Tables

List of Figures

List of Acronyms

Glossary

General Introduction 19

I Context and State of the Art 22

1 Industrial Context and the Fuel-Quality Problem 23

1.1 Cement manufacturing at Holcim El-Djazaïr – Oggaz 23

1.1.1 The dry-process route: preheater, calciner, rotary kiln 23

1.1.2 Thermal energy and the calciner’s two fuels 26

1.1.3 Alternative solid fuels and the Geocycle station 27

1.1.4 Fuel energy content as an unmeasured load on the thermal balance . . . 28

1.2 The fuel-quality problem 28

1.2.1 Heterogeneity and batch-to-batch variability of the ASF blend 28

1.2.2 The delayed, intermittent laboratory measurement 31

1.2.3 Operational consequences of unmeasured fuel quality 32

1.3 Available process instrumentation 33

1.3.1 What the plant measures 33

1.3.2 What the plant does not measure — and why this is an instrumentation
problem 33

1.4 Calorific value as an industrial variable 34

1.4.1 Laboratory calorific value (CV_lab) 34

1.4.2 Apparent process calorific value (CV_apparent) 35

1.4.3 Scope of the study 35

1.4.4 The online-estimation challenge and the question of observability . . . 35

1.5 Conclusion 36

2	State of the Art	37
2.1	Soft sensors	38
2.1.1	Physics-anchored and hybrid soft sensors	38
2.1.2	Estimating the calorific value of waste-derived fuels	39
2.2	Supervised learning for the empirical estimator	40
2.2.1	Gradient-boosted tree ensembles	40
2.2.2	Linear models and the bias-variance trade-off	41
2.3	Learning from aggregate labels	42
2.3.1	Multiple-instance learning as a precedent	42
2.3.2	Aggregate labels over time series	43
2.4	Digital twins and process simulation	43
2.4.1	Definition and levels of integration	43
2.4.2	Fidelity and applications	44
2.5	Multi-rate estimation and observability	44
2.5.1	Estimation under delayed and infrequent measurement	45
2.5.2	Observability and the masking effect of feedback	45
2.6	Validation under non-stationarity	46
2.7	Conclusion	46
II	Recovering the Signal	48
3	Data, Instrumentation and Methodology	49
3.1	The process and its instrumentation	49
3.1.1	Online process signals	50
3.1.2	The alternative-fuel injection system	51
3.1.3	What is measured online and what is not	52
3.1.4	The recorded signals: available, retained, and used	53
3.2	Laboratory measurement of calorific value	54
3.2.1	The measurement	54
3.2.2	The sample: grab sampling and the daily composite	55
3.3	Constructing the supervised dataset	56
3.3.1	Two records, one overlap	56
3.3.2	The composition of the label set	57
3.3.3	The dataset the modelling inherits	58
3.4	Validation methodology	58
3.4.1	Why random cross-validation does not apply	59
3.4.2	Out-of-sample evaluation on held-out later periods	60
3.4.3	A caveat the held-out scheme carries	60
3.5	The 2023 no-ASF reference period	60
4	From Steady-State Observability to the Disturbance Response	62

4.1	The steady-state channel	63
4.1.1	A reporting-scale artifact in the feed signal	64
4.1.2	The steady-state displacement is weak	65
4.2	Where the signal is not masked	67
4.3	The disturbance response	68
4.3.1	A single event: the feed cut of 13 May	68
4.3.2	Does it repeat? The event inventory	69
4.4	Constraints handed to the digital twin	71
4.4.1	Feed mass: quantised levels and a density bound	71
4.4.2	Moisture: the labels are on a net basis	72
4.4.3	Disturbance dynamics: the shape the twin must reproduce	72
4.5	Conclusion	73
5	Digital Twin of the Calciner	75
5.1	What the twin is, and what it is not	76
5.2	The forward model	77
5.2.1	Method and implementation	77
5.2.2	The nonlinearities	77
5.2.3	The controller	78
5.3	Constraints from the audit	78
5.4	Generating the labelled dataset	80
5.5	The seasonal calorific-value model	81
5.6	Does the twin behave like the real plant?	82
5.7	Conclusion	83
6	Recovering Calorific Value from the Twin	85
6.1	From labelled events to an estimator	85
6.1.1	The two model families	86
6.1.2	Two resolutions: single-event and daily-aggregated	87
6.2	Disturbance-response features	87
6.2.1	Shape features	87
6.2.2	The energy-balance feature	88
6.3	The two recovery routes	89
6.3.1	The energy-balance route	89
6.3.2	The shape route	89
6.4	Comparison, diagnosis, and selection	90
6.4.1	The full comparison across models and seeds	90
6.4.2	What governs recovery: the diagnostic sweep	92
6.4.3	Discriminating within a season, not merely across it	92
6.4.4	Selection	93
6.5	Conclusion	94

7	Early Daily Estimation	95
7.1	The operational question	95
7.2	What the estimate is built on: two timescales of fuel variation	96
7.3	Accuracy builds with observation time	96
7.4	What would sharpen the estimate: instrumentation	98
7.5	Conclusion	99
III	From Estimator to Instrument	100
8	Design of the Embedded Edge Gateway	101
8.1	System Architecture Overview	102
8.2	Hardware Design and Component Selection	103
8.2.1	Edge Computing Platform Selection	103
8.2.2	Proposed Power Supply Architecture	104
8.2.3	Proposed Bill of Materials	105
8.2.4	Mechanical Design Considerations	106
8.3	Software Pipeline and Event-Driven Inference	107
8.3.1	Design Rationale: Event-Driven over Fixed-Interval	107
8.3.2	Inference Workflow	107
8.3.3	The Linear Estimator	109
8.3.4	Daily Aggregation on the Gateway	110
8.4	Safety and Fault-Tolerance Mechanisms	110
8.4.1	Conformal Prediction Intervals	110
8.4.2	Fault-Tolerance and Watchdog Architecture	111
8.5	Cybersecurity and Network Architecture	111
8.6	Operator Dashboard Design	112
8.7	Performance Estimation and Feasibility Analysis	114
8.7.1	Theoretical Power Budget	114
8.7.2	Computational Timing and Real-Time Feasibility	114
8.7.3	Storage and Memory Margin	115
8.8	Laboratory Prototype Implementation and Validation	115
8.8.1	Prototype Hardware Configuration	115
8.8.2	LED Status Indicator System	117
8.8.3	Network Configuration and Dashboard	117
8.8.4	Prototype Experimental Observations	118
8.9	Integration with the Digital Twin	118
8.10	Conclusion	119
	General Conclusion	121
	What the work does not establish	122
	Perspectives and industrial developments	123

Closing	125
Bibliography	126

List of Tables

1.1	Representative components of the Oggaz ASF blend and their approximate net calorific value (MJ/kg).** Values are drawn from the Holcim El-Djazaïr – Oggaz laboratory characterisation of the co-processed waste streams.	29
3.1	Disposition of the principal recorded signals. Retained signals are those bearing on the calciner energy balance and on production-rate control; the rest are set aside as irrelevant, redundant, or administrative.	54
3.2	The 59 daily labels by provenance.	57
4.1	Calciner gas at matched steady-state production: the no-ASF 2023 reference against the two alternative-fuel periods.	65
4.2	Sustained feed-cut events, ordered by temperature-excursion depth. The two 13 May entries are distinct cut intervals on the same day, a short morning cut and a long afternoon cut. Only three events carry a fuel-CV label; the rest occurred on days without laboratory analysis.	69
4.3	The seven recurrent levels at which the alternative-fuel feed holds, with the fraction of active time spent at each.	71
4.4	Moisture coverage across the fuel-days, and the as-fired penalty where moisture is known.	72
6.1	Recovery across ten generating seeds: R^2 and Spearman rank correlation, mean $\pm$ standard deviation, for every route, resolution, and model.	91
6.2	The deciding comparison, condensed from Table 6.1: the shape route at daily resolution against the energy-balance route, with the mean absolute error added.	91
7.1	Accuracy of the daily calorific-value estimate against hours of the day observed, held-out seasons. Lead time is the interval before the end-of-day composite by which the estimate is available.	97
8.1	Industrial edge hardware evaluation matrix.	103
8.2	Proposed Bill of Materials (BOM).	106
8.3	Failure Modes and Effects Analysis (FMEA). All failure modes result in zero impact on production because the gateway operates in open-loop shadow mode with no write path to the DCS.	111

8.4	Estimated power budget (worst case, all components at maximum load).	114
8.5	GPIO-driven LED status indicator states on the prototype breadboard.	117

List of Figures

1.1	Schematic of the Oggaz dry-process line: preheater tower (C1–C5), calciner, rotary kiln and cooler, with the natural-gas and ASF feed points indicated at the calciner.	24
1.2	Panoramic view of the Oggaz plant from the Geocycle yard.	25
1.3	The Holcim El-Djazaïr – Oggaz process flow on the plant monitoring display: raw-meal preparation, preheater tower, calciner, rotary kiln and clinker handling.	25
1.4	The five-stage preheater and calciner tower at Holcim El-Djazaïr – Oggaz.	26
1.5	The rotary kiln.	27
1.6	Geocycle waste-composition board at the Oggaz plant, listing the co-processed waste streams and their approximate annual quantities.	29
1.7	Representative ASF feedstock streams. (a) ground olive pits; (b) acidic liquid waste; (c) tobacco waste; (d) expired pharmaceutical products; (e) insulation material and contaminated soil; (f) solid waste storage.	30
1.8	The Sludge Treatment Station. (a) An excavator homogenising the incoming sludge; (b,c) the storage basins from which the ASF stream is drawn.	31
1.9	What the plant measures and what it does not. Online signals capture the effects of fuel quality; the laboratory measures a chemical property of a sample, offline and delayed; the apparent calorific value, the cause the process responds to, has no online instrument. The soft sensor bridges the online signals to that missing quantity.	34
3.1	The Siemens CEMAT calciner control faceplate.	50
3.2	The Putzmeister injection-pump control screen at the Sludge Treatment Station.	51
3.3	The positive-displacement pump that injects the alternative-fuel stream into the calciner.	52
3.4	The alternative-fuel sampling station. At regular intervals a valve on the feed line opens and a grab sample is collected; the day’s grabs are combined into a composite, a portion of which is analysed in the laboratory.	55

3.5	The laboratory labels and the process historian on a common timeline. The laboratory labels (top) span late 2024 to mid-2026; the historian record (bottom) begins only in December 2025. Only the labels coinciding with active alternative-fuel feeding and falling within the historian period can be paired with online signals to form supervised examples; a label on a day without active ASF feed carries no disturbance to read and is excluded.	56
3.6	Distribution of the 59 labels by net calorific value, coloured by provenance (calorimeter and plant-simulated). The dashed line marks the mean and the shaded band one standard deviation; the dotted line marks the 3 MJ/kg physical floor below which five simulated recipe artifacts fall.	57
3.7	Monthly presence of each component in the alternative-fuel blend across the record. The composition shifts from month to month as the available waste streams change, so the fuel is non-stationary over the period.	59
4.1	Control architecture of the calciner temperature regulation loop.	63
4.2	Daily median of the non-zero alternative-fuel feed, logarithmic axis. The signal sits roughly eighteen times higher before 1 March 2026 than after, a hard reporting-scale step rather than a change in plant operation.	64
4.3	After dividing the pre-March feed by 18.2, the two reporting regimes overlay — both the gas-versus-feed scatter (left) and the feed distributions (right) coincide, confirming a single underlying process recorded on two scales.	65
4.4	Calciner gas at matched production for the no-ASF 2023 reference and the two alternative-fuel periods. The three means differ by less than the run-to-run variability of the gas: the alternative fuel’s displacement of gas is smaller than the noise floor, so the steady-state gas level cannot resolve it.	66
4.5	Daily matched-production gas against laboratory calorific value. A downward trend is present ($r = -0.53$ on this set) but loose, and it leans heavily on the single high-calorific-value point at the right.	67
4.6	The feed cut of 13 May 2026. When the alternative-fuel feed (top) is cut to zero, the calciner temperature (bottom) falls below its setpoint and the control system ramps the gas (middle) to recover it. Shaded bands mark the sustained ASF-off intervals.	68
4.7	Temperature-excursion depth against pre-cut feed level for the feed-cut events. Deeper excursions tend to follow higher pre-cut feed, as a fuel-energy interpretation predicts, though the relationship is loose and the labelled events are few. .	70
5.1	The digital twin as a one-directional calibrated forward model. Real plant data calibrate a forward model of the calciner, which generates labelled synthetic disturbance events for training the estimator. No path runs back to the plant: under the Kritzinger taxonomy this offline, one-way artifact is a digital model rather than a bidirectional digital twin.	76

5.2	A generated feed day. The feed snaps to discrete levels, holds each for two to three minutes, steps, and cuts to zero at intervals to create disturbance events, reproducing the quantised, held character of the real feed.	79
5.3	The twin's seasonal calorific-value profile, with monthly means set to the real plant values and a within-month batch spread. The composition-driven drift from roughly 6 to 18 MJ/kg across the year is reproduced explicitly.	81
5.4	Twin sanity checks over a simulated year. Left: the seasonal calorific-value drift is present. Right: across the year's cut events, excursion depth falls with fuel calorific value, a correlation of -0.48 , matching the loose, correctly-signed relationship of the real plant rather than an unrealistically clean one.	83
6.1	Left, the shape route's daily-aggregated predictions against true calorific value on held-out seasons. Right, recovery (R^2) by route and resolution: the shape route improves with aggregation while the energy-balance route does not, isolating feed-mass uncertainty as the cause.	90
7.1	The early daily-estimate accuracy as observation time increases. R^2 rises monotonically from 0.41 at one hour to 0.59 at seven hours, while the mean absolute error falls from 2.7 to 2.4 MJ/kg.	97
8.1	Open-loop cyber-physical architecture. The OT network (read-only) carries the Siemens CEMAT historian tags (ASF feed rate, C5 temperature, calciner gas, O_2 level at one-minute resolution) to the edge gateway (Raspberry Pi CM4) over a read-only OPC-UA link. The gateway performs event detection, extraction of the seven shape features, and evaluation of the linear estimator with a 90% conformal interval, then forwards only the CV estimate, confidence interval, and health status to the operator dashboard on the IT network. There is no write path to the DCS: the system operates strictly in open-loop shadow mode.	103
8.2	Power supply chain schematic: 24 V DC input through PTC fuse, TVS diode, EMI filter, and isolated DC/DC converter to the regulated 5 V rail supplying the CM4.	105
8.3	Gateway physical installation sketch: DIN-rail mounting inside the electrical cabinet adjacent to the CEMAT control panel, in a fanless IP54 enclosure. . . .	107
8.4	Event-driven inference pipeline flowchart: continuous OPC-UA acquisition and feed-cut detection in the outer loop, triggering event-window extraction, shape-feature computation, linear estimation, conformal interval, daily accumulation, and dashboard update in the inner loop.	109
8.5	Network segmentation architecture following IEC 62443 principles: physically separate ETH0 (OT, read-only OPC-UA) and ETH1 (IT, dashboard) interfaces with kernel-level IP forwarding disabled, so raw historian data never crosses the boundary.	112

- 8.6 Operator dashboard mock-up. Left panel: current daily CV estimate with conformal confidence interval, trend indicator, and event count. Right panel: thirty-day historical trend of soft-sensor estimates alongside available laboratory composite values. Bottom bar: real-time system health indicators. 113
- 8.7 Laboratory prototype photograph. The Raspberry Pi 4 Model B (2 GB RAM) is mounted on a breadboard with two GPIO-driven LED indicators (green, red) and connected to a laptop acting as the operator dashboard via a direct Ethernet cable. The inference pipeline, event-detection logic, and conformal prediction interval were all executed on this hardware during validation. 116

List of Acronyms

Acronym	Meaning
AI	Artificial Intelligence
ASF	Alternative Solid Fuel(s)
BOM	Bill of Materials
CM4	Raspberry Pi Compute Module 4
CEMAT	Siemens distributed control system (plant control system)
CV	Calorific Value
DCS	Distributed Control System
EMI	Electromagnetic Interference
FMEA	Failure Modes and Effects Analysis
GPIO	General-Purpose Input/Output
ID-fan	Induced-Draught fan
IP54	Enclosure protection rating: dust-protected and splash-resistant
IP5X	Dust-protected enclosure rating
IPX4	Splash-resistant enclosure rating
IT	Information Technology
MAE	Mean Absolute Error
MIL	Multiple-Instance Learning
NCV	Net Calorific Value (lower heating value)
OOP-UA	<i>(if you used this anywhere, correct to OPC-UA instead)</i>
OPC-UA	OPC Unified Architecture
OT	Operational Technology
PID	Proportional–Integral–Derivative (controller)
PTC	Positive Temperature Coefficient
TVS	Transient Voltage Suppressor
UPS	Uninterruptible Power Supply

Glossary

Term	Definition
Bomb calorimetry	The laboratory method by which a fuel sample is combusted in a sealed vessel to determine its calorific value, reported here on the net basis (NCV) in MJ/kg.
Calciner	The combustion chamber between the preheater and the rotary kiln in which the alternative fuel is burned and the bulk of the limestone decomposition takes place; its temperature is held at a setpoint by the control system.
Clinker	The nodular intermediate product of the kiln, formed when raw meal is heated to sintering temperature; it is ground with additives to make cement.
Geocycle	Holcim's waste-management and fuel-preparation operation, which sources, prepares, and doses the alternative solid fuels fed to the Oggaz plant.
Preheater	The multi-stage cyclone tower at the top of the kiln system, where the raw meal descends counter-current to the rising hot exhaust gases and is progressively heated before entering the calciner.
Raw meal	The finely ground feed to the kiln system, a blend of limestone, clay, and corrective minerals, fed to the top of the preheater tower.
Rotary kiln	The large inclined rotating cylinder downstream of the calciner, in which the calcined meal is sintered into clinker.
Sintering	The high-temperature stage, in the hottest part of the kiln, at which the calcined meal partially fuses and the clinker minerals form.

General Introduction

Cement is the binding agent of the built world, and making it is among the most carbon-intensive activities in industry. Cement production accounts for roughly seven percent of global anthropogenic carbon dioxide emissions. About forty percent of that share comes not from the chemistry of the material itself but from the fuel burned to reach the temperatures the process demands, where the kiln system must be held above 1400 degrees Celsius. Reducing the carbon intensity of cement therefore means, in large part, reducing the carbon intensity of the fuel, and the principal route open to plants today is substitution: replacing part of the fossil fuel with alternative solid fuels derived from waste.

These alternative fuels are attractive on two counts. They divert waste from landfill and incineration, and they displace coal, petroleum coke, and natural gas at a lower cost and a lower carbon burden. A modern plant can substitute a large fraction of its thermal energy this way. But the same property that makes these fuels cheap makes them difficult to burn well: they are heterogeneous. A stream blended from used oils, plastics, solvents, sludges, olive residue, and spent earth has an energy content that shifts from one delivery to the next and, at some plants, from one hour to the next. The calorific value of the fuel entering the kiln is not a constant to be looked up but a variable that moves.

That variability matters because the kiln is a tightly controlled thermal process. The temperature in the calciner, where the alternative fuel is burned, is held at a setpoint by a control system that trims the natural-gas supply up or down. When the alternative fuel runs rich, less gas is needed; when it runs lean, more is needed. If the controller knew the fuel's calorific value, it could anticipate these swings. It does not. At the plant studied in this work, the Holcim site at Oggaz in western Algeria, the only measurement of the fuel's calorific value is a laboratory analysis of a daily composite sample, available the following day. The control system runs on yesterday's number, or on no number at all, correcting for fuel-quality changes only after they have already disturbed the temperature.

This is the gap the present work addresses. The fuel's calorific value is operationally important, it changes faster than it is measured, and it is measured too late to act on. The natural question is whether it can be estimated instead from signals the plant already records continuously, the temperatures, gas flows, and combustion readings logged every minute by the control system. This is the idea of a soft sensor: an inferential estimate of a quantity that is hard to measure directly, built from quantities that are easy to measure.

But there is a reason to doubt that such an estimate is possible here, and it is the starting point of this thesis. The same control loop that makes the fuel's calorific value matter also works to hide it. A controller that successfully holds the temperature at setpoint does so by absorbing disturbances, and a change in fuel quality is, to the controller, a disturbance. If the loop does its job, the effect of a lean batch is cancelled by a gas adjustment before it shows as a temperature deviation, and the signal one would use to detect the lean batch is removed from the very measurements one would read it from. Whether anything of the fuel-quality signal survives this masking, and if so where, is not obvious and cannot be assumed.

Unlike much of the literature on data-driven soft sensors, which begins from the assumption that abundant and reliable labels exist and asks only which model predicts them best, this work begins one step earlier, by asking whether the quantity is observable at all under the instrumentation and data actually available in an operating cement plant. That difference in starting point shapes the entire thesis. It is not, in the first instance, a machine-learning exercise. It is an investigation into whether a physical quantity leaves a recoverable trace in a controlled process, and only then into how to recover it.

The investigation proceeds as a sequence of questions, each following from the answer to the last. The work is organised around six.

RQ1 — Observability. Is the apparent calorific value of the alternative fuel recoverable at all from the signals the plant already records, and if so, where in the process behaviour does the recoverable signal lie?

RQ2 — Digital twin. Where the real plant provides too few labelled examples to learn from, can a calibrated digital twin of the calciner be built to generate realistic, labelled disturbance events that the plant itself does not supply?

RQ3, Uncertainty. Can the principal physical uncertainties, in particular the fuel-mass uncertainty established from the feed measurement chain and the moisture uncertainty identified in the record, be quantified and built into that twin, so that the synthetic data it produces is honestly difficult rather than artificially clean?

RQ4 — Recovery. Given those generated events, can the calorific value be recovered, by which route, and how accurately on fuel compositions the estimator has not seen?

RQ5 — Limits of recovery. What physical factors fundamentally limit the recoverability of calorific value, and which uncertainties dominate the estimator's performance?

RQ6, Operational value. Can the resulting estimator be turned into something useful within a working day, an early estimate of the day's calorific value with useful lead time, and what would be needed to improve it?

RQ7 — Deployment. Can the estimator be turned into a practical embedded edge gateway for industrial use, under the constraints of low power, low computational load, read-only integration, and acceptable operator-facing output?

These seven questions trace the spine of the thesis, and each is taken up where the work can answer it. The remainder is organised as follows.

Chapter 1 establishes the industrial and technical setting: the dry-process kiln and its calciner, the alternative-fuel system at Oggaz, the heterogeneity of the fuel, and the instrumentation through which the process is seen, including the distinction between the laboratory calorific value and the apparent calorific value the process actually responds to.

Chapter 2 reviews the methods the work draws on: soft sensing, the gradient-boosted and linear model families and the trade-off that governs the choice between them, the weak supervision that arises when labels are aggregates over a day, the digital-twin taxonomy that lets the modelling artifact be named honestly, the control-theoretic notion of observability under feedback, and the validation methodology appropriate to a non-stationary process.

Chapter 3 describes the data: the online record and the laboratory record, how they overlap, how many usable labels result, what those labels are made of, and why the evaluation throughout holds out later periods rather than splitting at random.

Chapter 4 is the observability audit. It tests the steady-state channel and finds it weak, traces the weakness to the masking effect of the control loop, and locates the recoverable signal in the disturbance response, the transient that follows a cut in the fuel feed. It closes by quantifying the constraints the rest of the work must respect.

Chapter 5 builds the digital twin: a calibrated forward model of the calciner that reproduces the disturbance response under known fuel conditions, carries the measured mass and moisture uncertainties by construction, and generates the labelled events the real plant could not.

Chapter 6 trains and compares estimators on those events. It contrasts a route that reads the shape of the disturbance with one that reads the feed-based energy balance, identifies which succeeds and why, and uses the twin's controllability to diagnose what governs recovery.

Chapter 7 turns the estimator toward operation, asking how early within a day a usable estimate of that day's calorific value can be formed, and what instrumentation would sharpen it.

Chapter 8 designs and validates an embedded edge gateway for the estimator, first as a laboratory prototype and then as an industrial deployment blueprint.

The General Conclusion draws the answers together, states plainly what the work has shown and what it has not, and sets out what would be required to take the estimate from a calibrated result toward a working instrument.

Part I

Context and State of the Art

Chapter 1

Industrial Context and the Fuel-Quality Problem

The work reported in this thesis was carried out at the Holcim El-Djazair cement plant at Oggaz, where a substantial and growing share of the calciner's thermal demand is met by alternative solid fuels (ASF). These fuels are inexpensive and environmentally valuable, but their heat content is neither stable nor measured in time to be useful for control. This chapter establishes the industrial setting in which that problem arises. It first describes the process and the point at which the fuels are burned, and reframes the variable fuel quality as an unmeasured load that the control system must carry (§1.1). It then characterises the fuel-quality problem itself (§1.2), surveys the instrumentation the plant actually carries — and, by its absence, the one quantity it does not measure (§1.3). It then defines the variable this thesis sets out to estimate, distinguishing the laboratory quantity from the quantity the process actually experiences, and arrives at the question of observability (§1.4).

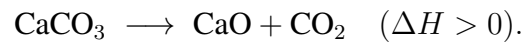
The guiding question is deliberately narrow: *why is this problem worth solving, and why is it hard?* The methods that answer it are reserved for the chapters that follow.

1.1 Cement manufacturing at Holcim El-Djazair – Oggaz

1.1.1 The dry-process route: preheater, calciner, rotary kiln

The Oggaz plant produces clinker by the modern dry process with a precalciner. Ground raw meal — a blend of limestone, clay and corrective minerals, is fed to the top of a multi-stage cyclone preheater tower, where it descends counter-current to the hot exhaust gases rising from below. Passing through the successive cyclone stages (denoted here C1 at the top down to C5 at the base, C5 being the hottest and the stage whose temperature the control system watches most closely), the meal is progressively heated before it reaches the **calciner**.

In the calciner, the meal undergoes calcination: the strongly endothermic decomposition of calcium carbonate into calcium oxide and carbon dioxide,



Calcination consumes the largest single fraction of the plant's thermal energy, and in a precal-ciner configuration it is deliberately moved out of the rotary kiln and into a dedicated combustion chamber so that the kiln itself can be shorter and the throughput higher. The calcined meal then enters the **rotary kiln**, an inclined rotating tube fired at its lower end, where the temperature rises to the sintering range and the clinker minerals form. The clinker is finally quenched in a cooler and conveyed to storage.

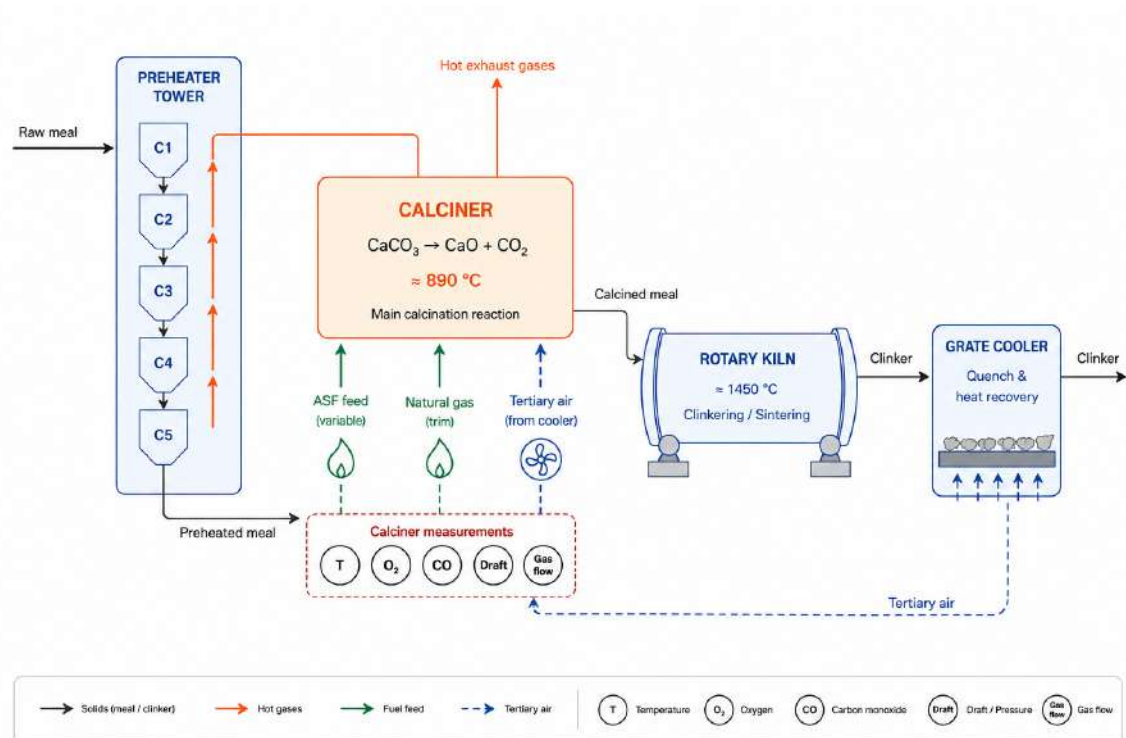


Figure 1.1. Schematic of the Oggaz dry-process line: preheater tower (C1–C5), calciner, rotary kiln and cooler, with the natural-gas and ASF feed points indicated at the calciner.



Figure 1.2. Panoramic view of the Oggaz plant from the Geocycle yard.

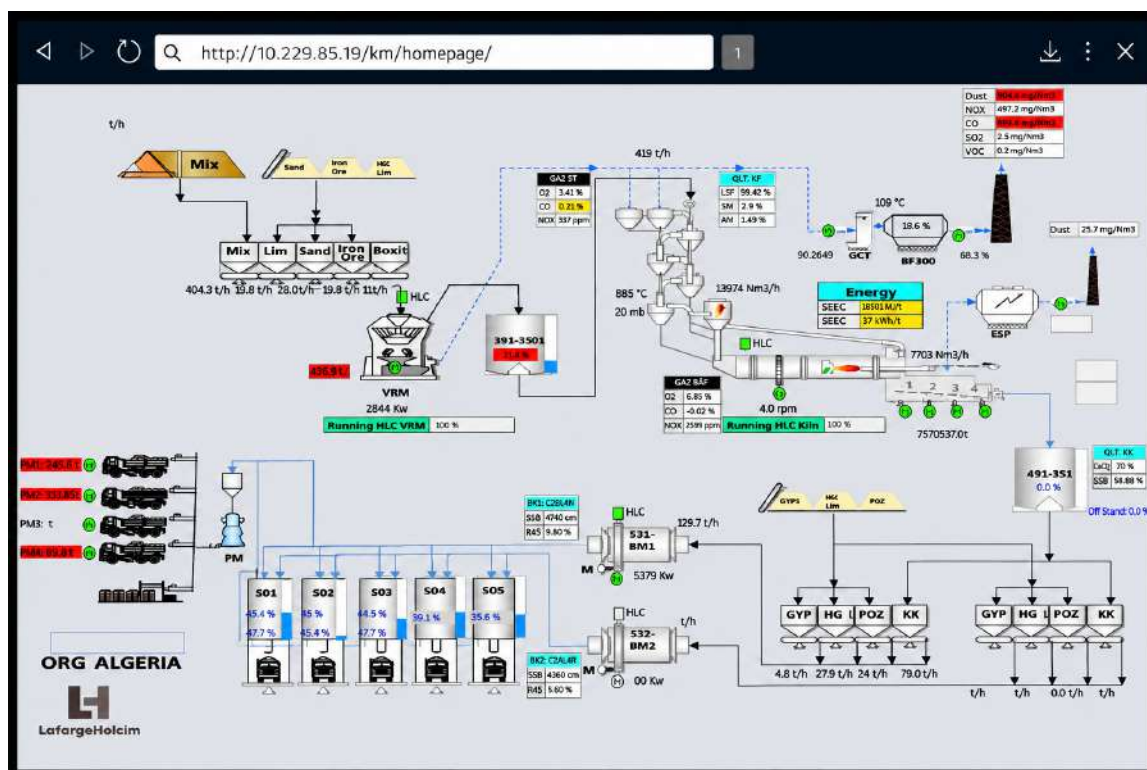


Figure 1.3. The Holcim El-Djazaïr – Oggaz process flow on the plant monitoring display: raw-meal preparation, preheater tower, calciner, rotary kiln and clinker handling.

The significance of this layout for the present work is that the calciner, not the kiln, is where the

alternative fuels are burned. It is therefore the calciner's measured signals that carry whatever information about fuel quality can be recovered.

1.1.2 Thermal energy and the calciner's two fuels

In a precalciner kiln the calciner typically accounts for the majority of the total fuel heat input — of the order of half to two-thirds — because calcination is so energy-intensive. At Oggaz this thermal demand is supplied by **two fuels burned together** in the calciner:

- **Natural gas**, the primary and baseline fuel. Natural gas is a clean, homogeneous fuel whose calorific value is essentially constant and well known (observed at approximately 36 MJ per normal cubic metre on the plant's gas readings over the study period; the value is approximate because gas composition can drift); its flow is metered directly, so its contribution to the calciner heat balance can be computed reliably. This treats the gas calorific value as a fixed reference; minor variations in gas composition and in combustion efficiency exist, and where the later analysis depends on this value the associated assumption is made explicit.
- **Alternative solid fuels (ASF)**, a prepared blend of waste-derived materials co-fired to displace part of the natural gas. Unlike the gas, the ASF stream is heterogeneous, variable, and, as the remainder of this chapter shows, neither directly metered for mass nor characterised for energy content at the moment it is fed.



Figure 1.4. The five-stage preheater and calciner tower at Holcim El-Djazair – Oggaz.

The control logic that ties the two together is central to everything that follows. A PID controller holds the C5 calciner temperature at its setpoint by trimming the natural-gas flow. The mechanism is reactive, not anticipatory: when the ASF releases less heat than before, the C5 temperature begins to drop; the controller detects this thermal error and raises the gas demand to compensate, and conversely backs the gas off when the ASF runs richer and the temperature tends to rise. The natural gas therefore acts as a continuously adjusted *balancing* fuel around the variable ASF contribution. Two points in this description are worth holding onto, because the thesis turns on them. First, the gas response is a *lagging* indicator: it follows a temperature error that the fuel has already caused, not a foreknowledge of the fuel. Second, this is precisely how a fuel that is never directly measured nevertheless leaves a measurable trace — in the gas the controller is forced to add or remove to keep the temperature flat.



Figure 1.5. The rotary kiln.

1.1.3 Alternative solid fuels and the Geocycle station

The ASF blend at Oggaz is prepared and dosed through Geocycle, Holcim’s co-processing operation. Co-processing is the practice of recovering energy from suitably prepared wastes by burning them in the cement kiln, where the high temperatures and long residence times destroy the organic fraction while the mineral fraction is incorporated into the clinker. It substitutes fossil fuel, diverts waste from landfill, and is a recognised route to lowering the net energy cost and carbon intensity of cement production (Kahawalage et al., 2023; Samolada and Zabaniotou,

2014) [1, 2]. The cement kiln is widely regarded as the most suitable industrial destination for solid recovered fuels, because its high temperatures and long residence times allow energy recovery while immobilising the mineral fraction in the clinker (Sarc et al., 2014) [3].

The fuels handled by the Geocycle station at Oggaz span a wide range of materials, among them used oils, expired industrial products, hydrocarbon-bearing sludges, spent bleaching earths, olive-pressing residues, and various low-grade solids. These are blended and pumped to the calciner. The degree to which a plant relies on such fuels is expressed by its thermal substitution rate (the share of thermal energy supplied by alternatives rather than fossil fuel); raising that rate is economically and environmentally desirable, but, as the next sections show, doing so safely requires knowing how much energy each batch of the blend actually carries.

1.1.4 Fuel energy content as an unmeasured load on the thermal balance

The control description in §1.1.2 can now be restated in the language an automation engineer would use, because doing so names the problem precisely. The variable an operator (or the feed controller) actually manipulates is the ASF **mass flow** — how many tonnes per hour of blend are pumped to the calciner. What governs the calciner’s heat balance, however, is not the mass flow but the **thermal input**, the product of that mass flow and the fuel’s energy per unit mass. Those two are equal only if the energy per unit mass is constant — and §1.2 will show that it is not.

The consequence is structural. Because the energy content of the blend shifts underneath a mass flow the operator believes is steady, *equal mass flow does not imply equal thermal input*. The calciner is therefore subject to an energy variation that no instrument reports: an **unmeasured load** on its heat balance. The control loop does not see this load directly; it sees only its eventual effect on the C5 temperature, and corrects after the fact. The plant’s thermal regulation is thus, in steady operation, a continuous compensation for an input it cannot measure. This single framing, fuel energy content as an unmeasured load on a closed loop — is what makes the problem a control-and-instrumentation problem before it is anything else, and it is the thread the later chapters follow when they ask where, if anywhere, that hidden load becomes visible in the signals the plant does record.

1.2 The fuel-quality problem

1.2.1 Heterogeneity and batch-to-batch variability of the ASF blend

The defining feature of the ASF stream is that it is assembled from components whose individual heat contents differ by more than an order of magnitude. This heterogeneity is intrinsic to waste-derived fuels: even standardised solid recovered fuels are classified into broad quality bands precisely because their calorific value, moisture, and composition vary widely between batches and sources (Sarc et al., 2014; Kahawalage et al., 2023) [3, 1]. Table 1.1 lists representative components of the Oggaz blend together with their approximate net calorific values.

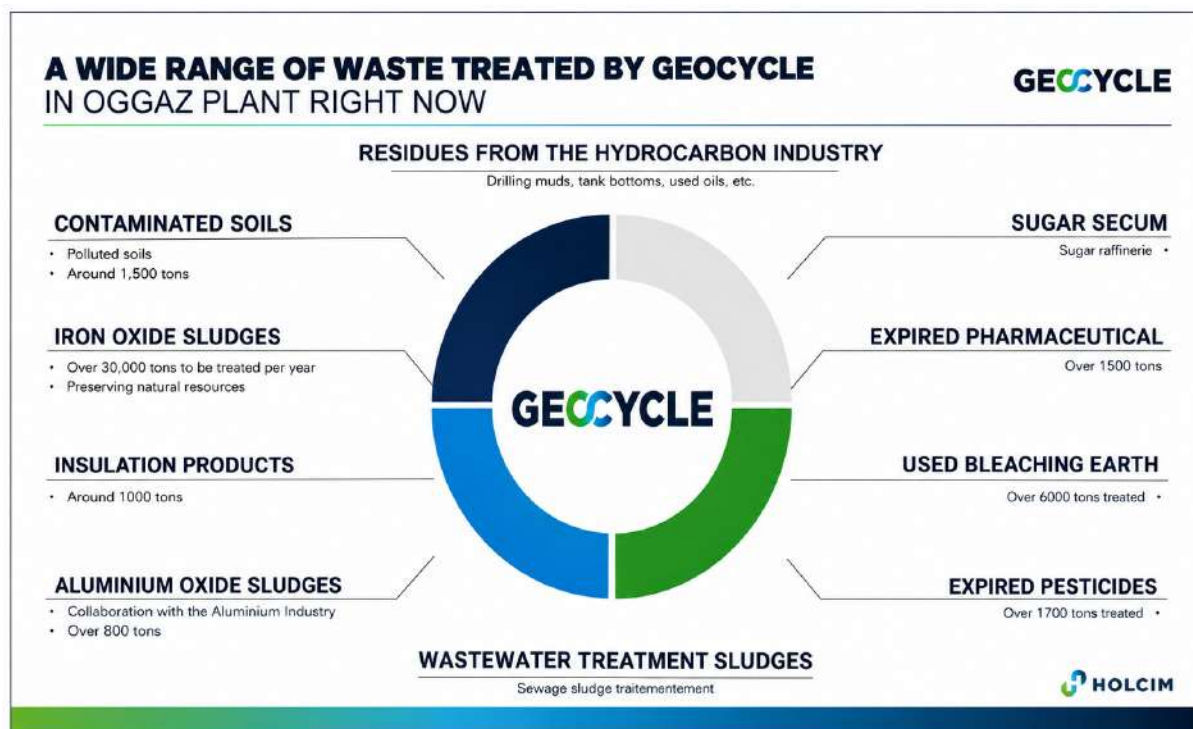


Figure 1.6. Geocycle waste-composition board at the Oggaz plant, listing the co-processed waste streams and their approximate annual quantities.

Table 1.1. Representative components of the Oggaz ASF blend and their approximate net calorific value (MJ/kg).** Values are drawn from the Holcim El-Djazair – Oggaz laboratory characterisation of the co-processed waste streams.

Component	Approx. CV (MJ/kg)
Used oil	45.8
Expired liquid fuel	36.1
Hydrocarbon sludge	33.8
Polyurethane	25.9
Glycol	17.0
Olive-pressing residue	13.5
Chemical product (misc.)	12.0
Spent bleaching earth	9.5
Filter-press (STEP) sludge	8.0
Sugar scum	6.0
Tobacco waste	5.6
Paint residue	5.4
Bentonite	2.0



Figure 1.7. Representative ASF feedstock streams. (a) ground olive pits; (b) acidic liquid waste; (c) tobacco waste; (d) expired pharmaceutical products; (e) insulation material and contaminated soil; (f) solid waste storage.

The observation to draw from Table 1.1 is the spread: a blend assembled from components ranging from roughly 2 to 46 MJ/kg cannot itself have a stable energy content. Its calorific value is whatever the day's recipe happens to make it. In practice the monthly mean of the delivered blend drifts across roughly 6 to 18 MJ/kg over the period studied (the seasonal profile is shown later, in Figure 5.3), and the day-to-day value moves within and around that drifting mean as the available waste streams change. Two consequences follow directly. First, the energy delivered to the calciner per kilogram of ASF is a moving target rather than a fixed fuel property. Second, and this is the harder consequence — because the recipe is driven by what waste is available rather than by a fixed specification, the variability has both a slow seasonal component (which months tend to be rich or lean) and a fast day-specific component (which particular load arrived today). The thesis returns to this decomposition in later chapters.



Figure 1.8. The Sludge Treatment Station. (a) An excavator homogenising the incoming sludge; (b,c) the storage basins from which the ASF stream is drawn.

1.2.2 The delayed, intermittent laboratory measurement

The plant does measure the calorific value of its fuels, but only in the laboratory and only after the fact. Fuel is sampled by grab sampling, and the samples are analysed by bomb calorimetry to yield a **Net Calorific Value** (NCV — the lower heating value), expressed in MJ/kg. The operationally trusted figure is the **24-hour composite**, a value representative of a whole day's fuel, which by construction is not available until the day is over, and in practice the following day.

This measurement has three limiting properties, and the third is the one that matters most for control. It is **intermittent**: a calorimetric value exists for only a fraction of operating days, and the remainder are covered, if at all, by recipe-based estimates rather than direct measurement. It is **coarse in time**: a single composite number stands for twenty-four hours of a fuel stream that, as §1.2.1 showed, varies within the day. And it is **delayed**, with a consequence that is easy

to understate. By the time the composite value for a given day is reported, the kiln has already burned that day's fuel, the C5 temperature has already moved in response, and the operators have already compensated with gas. The measurement therefore arrives after every decision it could have informed has been made. It can audit yesterday's energy accounting, but it is structurally incapable of supporting today's control. The laboratory answers "what was the fuel worth yesterday, on average?" A legitimate question for accounting, but not the operator's question, which is "what is the fuel worth now?"

1.2.3 Operational consequences of unmeasured fuel quality

Because the real-time energy content of the ASF is unknown, the plant operates the co-firing largely blind, and the consequences reach further than process stability alone.

The most immediate is **thermal stability**. An unexpected change in fuel quality propagates into the calciner temperature before any measurement can warn of it. When a lean or interrupted ASF load reaches the burner, the heat it was expected to supply does not arrive, the C5 temperature falls, and the controller must drive the gas up to recover.

This instability has knock-on costs. **Clinker quality** depends on stable calcination and a stable burning-zone temperature; thermal excursions of this size risk under or over burnt clinker and the mineralogical defects that follow. **Energy efficiency** suffers because the gas trim is necessarily reactive: the controller corrects after the temperature has already moved, so the gas is momentarily over or under supplied relative to what a foreknown fuel quality would have allowed; knowing the apparent fuel quality in advance would let the gas be *set* rather than *chased*. Reactive gas use and unstable combustion in turn raise **specific energy consumption and CO₂ emissions**, working directly against the environmental case that justifies co-processing in the first place. There is an **operator-workload** cost as well: lacking any forewarning of fuel quality, operators must watch the calciner closely and intervene by hand, carrying as vigilance what an instrument could carry as information.

Finally, the inability to see fuel quality **caps the substitution rate**. Pushing more ASF into the calciner is attractive economically and environmentally, but every increase in the alternative-fuel share also increases the plant's exposure to the blend's variability. Without a real-time estimate of what the blend is delivering, operators must keep a conservative margin precisely to protect against the batches they cannot see coming. A reliable estimate of apparent fuel quality is therefore not only a stability aid but an enabler of higher, safer substitution, the lever on which the economics and the emissions of the whole co-processing programme ultimately turn.

These consequences define the practical motivation for the thesis: the plant needs an estimate of fuel quality that is available *during* operation, from instrumentation it *already has*.

1.3 Available process instrumentation

Before asking why calorific value is not measured, it is worth establishing what *is* measured, because the plant is in fact heavily instrumented. The gap is specific, not general.

1.3.1 What the plant measures

Process signals at Oggaz are acquired through the plant's distributed control system and historian and are available at one-minute resolution. The recorded variables fall into four broad roles:

- **Thermal.** The C5 calciner outlet temperature (the controlled variable), the burning-zone temperature, the preheater-outlet temperature, and the secondary-air temperature.
- **Combustion and gas.** The calciner natural-gas flow and the total kiln gas flow, together with combustion-gas analysers reporting oxygen, carbon monoxide, nitric oxide and methane at the calciner, the kiln and the preheater exit.
- **Fuel and feed.** The ASF feed signal, the kiln raw-meal feed rate, and the clinker production rate.
- **Mechanical and draught.** The induced-draught fan speed, the kiln drive speed and torque, the primary-air pressure, and the C5 cyclone pressure.

This is a rich picture of the calciner's thermal and combustion state. The temperatures describe where the heat is, the gas flows and analysers describe how it is being produced and how completely the fuel is burning, and the mechanical and draught signals describe the aerodynamics that move the gases through the system.

1.3.2 What the plant does not measure — and why this is an instrumentation problem

Two absences in that list are decisive. First, there is **no online measurement of fuel energy content**: the only figure for ASF calorific value is the offline laboratory NCV described in §1.2.2, with all the limitations established there. Second, there is **no direct measurement of ASF mass flow**. The fuel is delivered by a pump, and the "feed" signal used by the plant is *derived* from the pump's displacement under an assumed material density rather than measured by a mass-flow meter — a point developed quantitatively in Chapter 3, but already enough to show that even the mass half of the thermal input is known only approximately. The plant thus measures the *effects* of fuel quality on the process in great detail while measuring the *cause* — the energy actually entering with the fuel, not at all.

From an engineering perspective, this work is situated in the field of instrumentation and virtual sensing. Industrial measurements acquired through the plant historian are processed to estimate a quantity that cannot be measured continuously, thereby forming a **virtual (inferential) instrument**. The resulting estimator is ultimately implemented on **embedded edge hardware** for

real-time operation, linking industrial sensing, signal processing, soft sensing, and embedded systems within a single instrumentation framework. Figure 1.9 summarises this measurement landscape: the quantities the plant measures online, the calorific value it measures only offline in the laboratory, and the apparent calorific value it does not measure at all, which the soft sensor must infer from the online signals.

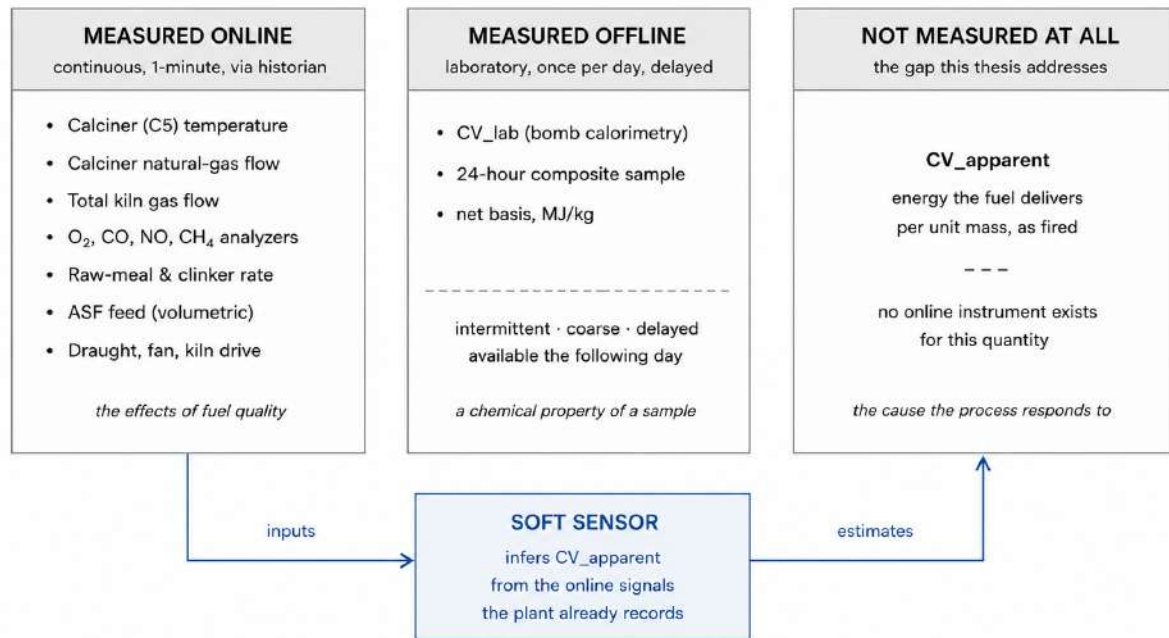


Figure 1.9. What the plant measures and what it does not. Online signals capture the effects of fuel quality; the laboratory measures a chemical property of a sample, offline and delayed; the apparent calorific value, the cause the process responds to, has no online instrument. The soft sensor bridges the online signals to that missing quantity.

1.4 Calorific value as an industrial variable

The variable at the centre of this work is the calorific value (CV) of the alternative fuel. It is essential to distinguish two distinct quantities that share that name, because the thesis is about one of them and not the other, and the difference between them is itself a result.

1.4.1 Laboratory calorific value (CV_lab)

The laboratory calorific value, denoted **CV_lab**, is the intrinsic energy content of a fuel sample as determined by bomb calorimetry, expressed on the net basis (NCV) in MJ/kg. It is a property of the chemistry of the sample: it depends on what the fuel is made of, not on how or where it is burned. CV_lab is the quantity reported by the plant laboratory and tabulated by component in Table 1.1. It is accurate for the sample analysed, but, as established in §1.2.2, it is delayed, intermittent, and tied to a grab sample that may not represent the stream actually entering the calciner over the following hours.

1.4.2 Apparent process calorific value (CV_{apparent})

The apparent process calorific value, denoted **CV_{apparent}**, is a different quantity: it is the energy the fuel actually delivers to the calciner per unit mass fed, as the process experiences it rather than as a laboratory measures it. Where CV_{lab} is a chemical property of a sample, CV_{apparent} is a process property of the fuel as fired. It is the quantity that matters operationally, because it is the one that governs how the calciner responds and therefore the gas the controller must supply. It is also the quantity this thesis sets out to estimate; the reasons it can differ from the laboratory value are not assumed here but established later, from the data.

1.4.3 Scope of the study

The aim of this work is not to add physical instrumentation to the production line, but to determine how far the information already recorded by the plant's control system can be used to estimate a quantity that line does not measure. The study is therefore bounded by what the existing sensors provide: it works with the signals defined in the previous section, under the operating and measurement limitations those sensors impose, and asks what can and cannot be recovered from them.

The chapters that follow develop this investigation in order — first the available data and instrumentation and their preparation, then the methodology proposed to estimate the calorific value, then its validation, and finally its realisation as a proof-of-concept virtual instrument.

1.4.4 The online-estimation challenge and the question of observability

Putting §1.2 to §1.4 together defines the technical challenge precisely. There is no instrument at Oggaz that measures fuel calorific value online, so any real-time estimate must be inferred from the signals the plant already records, principally the calciner gas demand, the C5 temperature, and the combustion-gas readings, through a soft sensor. Whether those signals carry enough information to support such an estimate is not assumed in advance; it is the question the investigation has to settle.

This is what raises the matter from an estimation task to a question of **observability**. The central question, then, is not whether the alternative fuel has a calorific value — it plainly does, but whether that value is *observable* from the instrumentation the plant already carries: whether the unmeasured load of §1.1.4 leaves, somewhere in the recorded signals, a trace from which it can be recovered, and if so, at what time resolution and with what accuracy. Whether a useful estimate can be recovered at all under these conditions is an open question rather than a foregone conclusion, and it is the question this thesis sets out to answer honestly.

The preceding sections have shown why the calorific value of the alternative fuel is operationally important, why it cannot be measured continuously with the available instrumentation, and why the existing online variables provide only indirect evidence of it. Whether that apparent calorific value can nevertheless be recovered from those variables, and under what limitations, is the

question taken up in the chapters that follow.

1.5 Conclusion

This chapter has set out the industrial problem the thesis addresses. The Oggaz plant co-fires a heterogeneous blend of alternative solid fuels in its calciner, where they supply a large part of the thermal demand. Because that blend is assembled from components differing in heat content by more than an order of magnitude, its energy per unit mass varies from day to day, while the mass flow the operator manipulates stays under his control — so the calciner is subject to an *unmeasured load* on its heat balance, which the control loop can only compensate after the fact. The plant measures the effects of this load on temperature and gas in detail, yet measures neither the fuel's energy content nor, directly, its mass.

The variable of operational interest is consequently not the laboratory calorific value but the apparent calorific value, the energy the fuel delivers to the process per unit mass. Naming that target is the work of this chapter; establishing how, and how well, it can be recovered from the existing signals is the work of those that follow. The absence of any online instrument for this variable carries concrete costs in stability, clinker quality, energy, emissions, operator workload, and the achievable substitution rate. Whether the apparent calorific value can be recovered from the existing signals, and under what limitations, is the question taken up in the chapters that follow.

Chapter 2

State of the Art

The previous chapter established the problem: the apparent calorific value of the alternative fuel stream at Oggaz is an operationally important quantity that is measured only once per day, offline, from a composite sample, while the process it drives is monitored continuously. Estimating that quantity from the available online signals is the task this thesis takes up. It is not a new kind of task in the abstract — estimating an expensive or slow measurement from cheap continuous ones is a recognised engineering problem with an established body of methods, and the purpose of this chapter is to locate the present work within that body, identify the techniques it draws on, and make clear where the standard tools apply cleanly and where the particular features of this problem will require care.

The chapter is organised around the components the later design will assemble. Section 2.1 introduces soft sensors, the general class of estimators to which the proposed instrument belongs, the distinction between data-driven and physics-anchored designs, and the prior work on estimating fuel calorific value against which this thesis is positioned. Section 2.2 reviews the supervised learning methods used for the empirical part of the estimator — the gradient-boosting family of tree models and the linear baselines against which they are compared — together with the bias-variance and model-parsimony considerations that govern which is appropriate when data is limited. Section 2.3 covers the problem of learning from aggregate labels, where a single daily value supervises a sequence of higher-rate observations. Section 2.4 turns to the digital twin, the physics-based component of the design, and uses the established taxonomy of the field to position the artifact built here honestly within it. Section 2.5 reviews the multi-rate and observability literature that frames estimation of a slowly, irregularly measured variable in a controlled process. A final section draws together the validation methodology appropriate to a non-stationary process of this kind.

Throughout, the aim is not an exhaustive survey but a targeted one: each method is reviewed to the depth needed to justify its use, or its rejection, in the chapters that follow.

2.1 Soft sensors

A soft sensor (or virtual sensor) is a predictive model that estimates a process variable which is difficult, expensive, or slow to measure directly, using other variables that *are* measured continuously and cheaply. The name contrasts these software-based estimators with the “hard” physical instruments they stand in for. In the process industries they are used wherever a quantity of interest is available only through laboratory analysis, an offline assay, or an instrument too costly to deploy on every stream, while a set of correlated online measurements, temperatures, pressures, flow rates, power draws — is recorded continuously by the plant control system (Fortuna et al., 2007) [4]. The soft sensor learns or encodes the relationship between those online measurements and the target quantity, and then supplies an estimate at the rate of the online signals rather than the rate of the laboratory.

Two broad families of soft sensor are usually distinguished by how that relationship is obtained. *Model-driven* (or first-principles) soft sensors derive the relationship from the governing physics and chemistry of the process — mass and energy balances, reaction kinetics, thermodynamic relations, so that the estimate is grounded in mechanism and remains interpretable in engineering terms. *Data-driven* soft sensors instead infer the relationship empirically from historical records of the inputs and the target, using statistical or machine-learning regression, and make no explicit commitment to the underlying mechanism (Jiang et al., 2021) [5]. Each family carries a characteristic trade-off: first-principles models extrapolate more reliably and explain their estimates, but require a process well enough understood to be written down and calibrated; data-driven models are faster to build and can capture relationships too complex to model explicitly, but depend on the quantity and representativeness of the available data and can fail without warning when the process moves outside the conditions it was trained on (Jiang et al., 2021; Fortuna et al., 2007) [5, 4].

2.1.1 Physics-anchored and hybrid soft sensors

Because the two families fail in different ways, a substantial line of work seeks to combine them, anchoring data-driven estimators in physical structure so that the empirical model is constrained by what is known about the process rather than learning the relationship from scratch. Such hybrid or physics-anchored soft sensors use first-principles knowledge to shape the model — for example by supplying physically meaningful inputs, by enforcing known balances as constraints, or by using a mechanistic model to generate or correct the data the empirical model learns from — and report improved robustness and data efficiency relative to purely data-driven designs, particularly when measured data is limited or unevenly distributed across operating conditions (Liu et al., 2024) [6]. This combination of a physical description of the process with an empirically trained estimator is the strategy on which the present work builds, and the relevant components, the gradient-boosting and linear regression methods used for the empirical part, the digital-twin concept used for the physical part, and the multi-rate estimation setting that frames the whole problem — are reviewed in the sections that follow.

2.1.2 Estimating the calorific value of waste-derived fuels

The specific target of this thesis — the calorific value of a waste-derived fuel, has itself been the subject of estimation work, and it is worth being precise about what has been attempted, because it locates the contribution made here. The established approach estimates calorific value from the fuel's *composition*: empirical and machine-learning models predict the heating value of refuse-derived and solid recovered fuels from their ultimate or proximate analysis, the elemental or component breakdown obtained by laboratory assay (Mateus et al., 2021; Tahir et al., 2022) [7, 8]. These models are accurate and useful, but they are offline by nature: they require a chemical analysis of a fuel sample, and so answer the question "given this analysed sample, what is its calorific value?" rather than the operational question of what the fuel entering the process is delivering now.

A second, smaller line of work estimates heat content *online*, from process data rather than from sample composition. In waste incineration, Lim et al. (2024) [9] recover the heating value and composition of the burned waste from plant operating data using an energy-balance reverse-calculation, explicitly to overcome the absence of real-time information about a variable fuel. That setting is the closest precedent to the present one in spirit — inferring a latent fuel-quality variable from the process it drives — but it differs in two decisive respects: it is a dedicated waste incinerator rather than a cement calciner co-fired with a balancing fuel under closed-loop control, and it relies on additional instrumentation installed for the purpose, whereas the present work is restricted to the signals the plant already records. The cement-specific literature on alternative fuels, by contrast, has concentrated on combustion behaviour and process implications, for example computational studies of solid-recovered-fuel co-firing inside a calciner (Mikulčić et al., 2016) [10] — and on the offline characterisation and quality classification of the fuels themselves (Sarc et al., 2014) [3], rather than on online recovery of their calorific value. In commercial practice, AI-based kiln-optimisation systems predict *downstream* clinker-quality indicators such as free lime and re-optimize setpoints in response, treating fuel variability as a disturbance to be absorbed rather than a quantity to be measured.

The position of this thesis follows from that survey. Predicting calorific value from fuel composition is established but offline; recovering heat content online from process data has been demonstrated in incinerators, with added instrumentation and via an energy balance; and the cement literature addresses co-firing behaviour and offline fuel quality rather than online calorific-value recovery. What has not, to the author's knowledge, been directly addressed is the online recovery of the *apparent* calorific value of an alternative fuel in a cement calciner operating under closed-loop control, using only the instrumentation the plant already carries — the problem this thesis takes up. The energy-balance route of the incineration work reappears here as one of the two recovery routes compared in Chapter 6, where its sensitivity to feed-mass uncertainty is examined directly.

2.2 Supervised learning for the empirical estimator

The empirical part of a data-driven soft sensor is a supervised regression problem: given a set of online measurements as inputs and recorded values of the target as labels, learn a function that maps the former to the latter and generalises to new data. Two model classes are relevant to the present work and are reviewed here, gradient-boosted tree ensembles, which can capture complex nonlinear relationships, and linear models, which serve as the interpretable baseline against which the ensembles are judged. The question of which class is appropriate for a given problem is not settled in the abstract; it depends on properties of the data and the underlying relationship, and the second half of this section reviews the principle that governs the trade-off.

2.2.1 Gradient-boosted tree ensembles

Gradient boosting builds a predictive model as an additive ensemble of simple base learners, almost always shallow regression trees, fitted in sequence. Each new tree is trained to correct the errors of the ensemble built so far, by fitting the negative gradient of a chosen loss function — a procedure Friedman (2001) [11] formalised as gradient boosting and which generalises boosting from classification to arbitrary differentiable losses, including the squared error used in regression. The result is a flexible nonparametric estimator that captures nonlinear relationships and interactions between inputs without their being specified in advance, while the use of shallow trees and a learning rate controls how quickly the ensemble fits the data. Formally, the ensemble is built stagewise as an additive expansion,

$$F_M(x) = \sum_{m=1}^M \nu h_m(x) \quad (2.1)$$

where each base learner h_m is a shallow regression tree, ν is the learning rate, and M the number of stages. At each stage m , the new tree is fitted to the negative gradient of the loss evaluated at the current ensemble; for the squared-error loss used in regression, this negative gradient is simply the residual,

$$r_{i,m} = - \left[\frac{\partial \frac{1}{2} (y_i - F(x_i))^2}{\partial F(x_i)} \right]_{F=F_{m-1}} = y_i - F_{m-1}(x_i) \quad (2.2)$$

so each successive tree corrects the errors left by its predecessors.

Several widely used implementations refine this idea for scale and speed. XGBoost (Chen and Guestrin, 2016) [12] adds regularisation and an efficient sparsity-aware split-finding algorithm, and has become a standard tool in applied regression and classification. LightGBM (Ke et al., 2017) [13] introduces histogram-based binning of feature values and a leaf-wise growth strategy, substantially reducing training time and memory on large datasets; the histogram-based gradient-boosting estimators now available in general-purpose libraries derive from this line of

work. These methods are accessible through standard open-source toolkits, of which scikit-learn (Pedregosa et al., 2011) [14] is the most widely used in Python and provides both the gradient-boosting estimators and the linear baselines discussed below.

Gradient-boosted trees have an established record on industrial process-modelling problems, including in the cement industry. Fatahi et al. (2022) [15] applied gradient-boosted trees to model energy-consumption factors of an industrial cement vertical roller mill, reporting predictive accuracy above $R^2 = 0.80$ on held-out data and outperformance of conventional baselines such as random forests and support-vector regression. That study is a useful methodological precedent: it demonstrates that gradient-boosted trees can recover a useful predictive relationship for a process variable in a cement plant from routinely logged operating signals. The target there differs in nature from the one addressed here, however — Fatahi et al. model a mechanical power consumption that is itself continuously and directly measurable, whereas the present work estimates a latent fuel-quality variable available only through intermittent laboratory assay. What transfers from that work is therefore the modelling apparatus, not the structure of the estimation problem, which is considerably harder here because the target itself is scarce and indirectly observed.

2.2.2 Linear models and the bias-variance trade-off

Against these flexible ensembles stands the linear model, regression in the original inputs or in a small set of engineered features — which is the natural baseline in any regression study. A linear model has few parameters, is interpretable in that each coefficient has a direct reading, and is cheap to fit and to deploy. Its limitation is equally clear: it can represent only linear relationships, and will underfit a process whose true input-output relationship is strongly nonlinear.

The choice between a flexible ensemble and a simple linear model is an instance of the bias-variance trade-off (Hastie, Tibshirani and Friedman, 2009). A flexible model has low bias — it can represent complex relationships, but high variance: with limited or noisy training data it fits the noise as well as the signal, and generalises poorly. A simple model has the opposite profile: higher bias, but lower variance and more stable behaviour on small samples. Which profile is preferable is not a matter of principle but of regime, and the relevant factor is how well the complexity of the model matches the complexity of the underlying process relative to the amount of data available.

Recent work in statistical learning theory makes this conditional precise. Bargagli Stoffi, Cevolani and Gnecco (2022) [16] show, within the framework of structural risk minimisation, that a preference for simpler models provably aids learning when the data-generating process is itself simple and the training set is small, but can hinder learning when the process is genuinely complex and data is scarce — so that simplicity is helpful exactly when the underlying relationship is simple, not as a universal rule. Dubova et al. (2025) [17] reinforce this caution from the opposite direction, arguing across modern modelling practice that parsimony is not a universal guide to predictive accuracy and that complex models are sometimes essential. Taken together, this literature frames model selection as regime-dependent: the appropriate complexity of the esti-

mator is a property to be determined for the specific problem and data, not assumed in advance. The considerations discussed above provide the framework within which the empirical models examined later in this thesis are selected and evaluated.

The effective complexity of a regression problem depends not only on the model class but also on the representation of the input variables. Carefully engineered features can simplify relationships that would otherwise require more flexible nonlinear models.

2.3 Learning from aggregate labels

The supervised methods reviewed above assume that each input observation carries its own label. The present problem departs from that assumption in a way that shapes the whole estimation task: the target is measured once per day, from a composite sample, while the inputs are recorded many times within that same day. A single daily calorific value therefore supervises a whole sequence of higher-rate input observations, with no label available for any individual observation inside the day.

This is a problem of *aggregate-label weak supervision*: the available label is a single value attached to a collection of observations, and it is an aggregate over that collection — here, the calorific value of a composite formed by mixing the day’s grab samples, rather than a property of any one observation within it. The estimation task must work with supervision at the level of the daily aggregate while the inputs, and ultimately the desired estimates, live at a finer time scale. One consequence of this scheme deserves to be flagged at the outset: attaching a single daily value to every window within the day introduces a known source of label noise where the fuel quality varies within the day, since each window is supervised by the daily average rather than its own true value. This is a structural limitation of aggregate-label supervision, returned to when the labelling scheme is constructed and again when the within-day variability is examined.

2.3.1 Multiple-instance learning as a precedent

One of the best-established machine-learning frameworks for labels attached to collections rather than to individual observations is multiple-instance learning (MIL), introduced by Dietterich, Lathrop and Lozano-Pérez (1997) [18]. In MIL, training labels are attached to *bags* of instances rather than to the instances themselves — originally for a drug-activity problem where a molecule’s activity was known but the responsible conformation within it was not — and the learner must infer instance-level structure from bag-level labels alone.

MIL and the present problem share the defining feature, a label that applies to a collection rather than to its members, but they differ in how that label relates to the collection. Classic MIL uses a maximum-style rule, in which a single triggering instance fixes the bag label; the daily composite here is instead an *average* over the day’s fuel, an aggregate of all the instances rather than a property of one. MIL is therefore the right conceptual precedent for the situation, supervision at the level of a collection — but not a direct method for it, and the distinction between a triggered

and an averaged label is one the thesis returns to when the labelling scheme is constructed.

2.3.2 Aggregate labels over time series

Because a sequence of timestamped observations is naturally a collection of instances ordered in time, the aggregate-label setting has been studied for time series. Chen et al. (2024) [19] treat multivariate time-series classification as a weakly supervised problem in which each time point is an instance and an entire series is a bag, on the observation that the patterns relevant to a series-level label are often sparse and localised within a long sequence rather than present at every moment. Their method preserves the temporal ordering and dependence between instances rather than treating them as exchangeable.

This is a useful precedent for supervising a sequence of process measurements with a single collection-level label, and for the idea that the informative content of a labelled window may be concentrated in a few of its moments rather than spread evenly across it. It remains, however, a classification method, whereas estimating a continuous calorific value is a regression problem, and the daily label here is an aggregate rather than a sparse trigger. The present work therefore takes from this literature the conceptual framing rather than the specific learning algorithm.

2.4 Digital twins and process simulation

The physics-anchored strategy introduced in Section 2.1 requires a model of the process that can stand in for it — one that reproduces how the measured signals respond to changes in operating conditions, so that the behaviour of the real plant can be studied, and data generated, without perturbing production. In current usage such a model is most often called a *digital twin*. The term is now widespread but loosely applied, and using it precisely requires the distinctions the field has developed.

In the present work, the process model serves two purposes: it provides a controlled environment in which the response of the measured signals to fuel-quality variation can be studied, and it generates physically consistent operating scenarios that complement the limited labelled plant data. Both purposes are taken up in the later chapters; the present section establishes the concept and the vocabulary they rely on.

2.4.1 Definition and levels of integration

The most widely cited framework for this purpose is that of Kritzing et al. (2018) [20], who classify digital representations of a physical system by the degree of data integration between the physical and the digital object. They distinguish three levels. A *digital model* is a digital representation with no automated data exchange: a change in the physical object does not automatically update the digital one, and vice versa. A *digital shadow* adds an automated one-way flow, from the physical object to the digital one, so that the digital representation tracks the state of the physical system but cannot act back on it. A *digital twin*, in the strict sense, has automated

data flow in *both* directions, so that the digital object not only mirrors the physical system but can also feed information or control back to it. Kritzinger et al. note that much of the literature uses the word "digital twin" for what are, by this criterion, digital models or shadows, since true bidirectional integration is comparatively rare.

This taxonomy is the one adopted in the present work, because it allows the artifact built here to be positioned precisely. Throughout this thesis, the term *digital twin* is used in the broader industrial sense commonly found in the process-engineering literature, while recognising that, under the stricter taxonomy of Kritzinger et al. (2018) [20], the developed artefact corresponds more closely to a *digital model*, or to a digital shadow during the replay of historical data — since it is calibrated offline and does not exchange data with the plant in a closed loop during operation. Stating its level of integration here ensures the term is not read as a stronger claim than the work supports.

2.4.2 Fidelity and applications

Two further strands of the digital-twin literature bear on this work. The first concerns *fidelity* — how closely the digital representation must reproduce the physical system, and in what respects. Jones et al. (2020) [21], in a systematic review, identify fidelity as one of the defining and still under-specified characteristics of a digital twin, noting that a representation need not reproduce every aspect of the physical system but only those relevant to its intended use. This is directly pertinent here, since the process model must reproduce the process behaviours relevant to the estimation task, rather than every aspect of the physical plant, and the question of which behaviours it must capture, and to what accuracy, is one the design chapters address explicitly. The second strand concerns scope of application: Tao et al. (2019) [22] survey the use of digital twins across industrial functions including design, production, and prognostics, establishing the breadth of the concept and locating process-level simulation of the kind used here within it.

Taken together, this literature supplies both the vocabulary to describe the physical component of the soft sensor accurately and the criteria, level of integration, and fidelity relative to purpose — against which that component is characterised in the chapters that follow.

2.5 Multi-rate estimation and observability

The components reviewed so far — the soft-sensor framing, the empirical models, the aggregate-label supervision, and the process model, describe *how* an estimate might be built. This section addresses a prior question: *whether* the target is observable from the available signals at all, and what makes that recovery difficult. The relevant framework is the control-theoretic notion of observability, together with the literature on estimation when measurements arrive at a slower and more irregular rate than the process evolves.

2.5.1 Estimation under delayed and infrequent measurement

The defining feature of the present problem is a mismatch of rates: the process and its online signals evolve continuously, while the target is measured once per day, from a sample integrated over that day, and with the laboratory result available only after a delay. This is a multi-rate estimation problem, and it is well studied in process control. Guo and Huang (2015) [23] address state estimation when a measurement is both delayed and *integral* — that is, when the available reading reflects an accumulation over an interval rather than an instantaneous value — which is precisely the character of a daily composite assay. Their work shows that such measurements can still inform estimation, but that doing so requires explicitly accounting for the integrating, delayed relationship between the measurement and the underlying state, rather than treating the reading as a direct sample.

A related strand concerns how confidence in an estimate should be represented when measurements are sparse. Dürr et al. (2021) [24] develop a cascaded estimation scheme in which the uncertainty of an estimate is carried explicitly as a variance that grows between measurements and contracts when a new measurement arrives. This treatment of confidence as a time-varying quantity is relevant to any soft sensor that must produce estimates continuously while being corrected only intermittently, as is the case here.

2.5.2 Observability and the masking effect of feedback

Whether a quantity can be reconstructed from a given set of measurements is, in control terms, a question of observability. For the present problem this question is sharpened by a feature of the plant: the process is under closed-loop control, with the manipulated inputs adjusted automatically in response to the very disturbances whose source one would like to estimate. A closely analogous difficulty is long recognised in the related field of closed-loop system identification. Ljung (1999) [25] and the foundational analysis of Gustavsson, Ljung and Söderström (1974) [26] show that feedback can mask the information needed to identify a process, because the controller's action removes from the measured signals part of the variation that would otherwise reveal the underlying relationship. While those results concern the identification of a process model rather than the estimation of a latent input, the mechanism they describe carries over by analogy: a well-regulated loop, by holding a controlled variable near its setpoint, can suppress exactly the signal excursions that would carry information about an upstream disturbance.

This has a direct bearing on the estimation of fuel quality in a controlled process. If the control system compensates for variations in fuel energy by adjusting another input, the effect of those variations may be partly absorbed by the loop rather than appearing as a free excursion in the measured signals. Whether, and under what conditions, the fuel-quality signal remains observable in the available measurements is therefore not a question that can be settled in the abstract; it depends on the specific structure of the loop, the measurements available, and the conditions under which the plant is operating. Establishing this for the Oggaz process is one of the cen-

tral tasks of the chapters that follow, and the present review serves to frame it as a question of observability under closed-loop control rather than as a simple matter of fitting a model to data.

2.6 Validation under non-stationarity

A soft sensor is only as trustworthy as the procedure used to estimate its error, and the choice of that procedure is itself a methodological question with a literature. The default in machine learning is k-fold cross-validation, in which the data are randomly partitioned into folds and each fold is held out in turn. This default rests on an assumption, that observations are independent and identically distributed — which fails for time series, where observations close in time are dependent and the process may evolve over the period of record.

Bergmeir and Benítez (2012) [27] set out the consequences of this for time-series model evaluation, showing that the i.i.d. assumption underlying standard cross-validation is invalidated by temporal dependence, and that reserving a contiguous block from the end of the series for testing removes the theoretical difficulty by respecting the temporal order. Cerqueira, Torgo and Mozetič [28] extend this into an empirical comparison across many real-world series, with a result that bears directly on the present work:: for *stationary* series, blocked cross-validation can be applied, but for *non-stationary* series the most accurate error estimates are produced by out-of-sample evaluation, in which the model is trained on an earlier period and tested on a later one. Because the fuel-quality process here is non-stationary by construction — its composition drifts as the available waste streams change through the year, this literature recommends evaluating the estimator on held-out later periods rather than by random cross-validation. The validation strategy used in this thesis follows that recommendation.

This choice carries a caveat that the literature is equally clear about. Roberts et al. (2017) [29] show that while blocking the data to respect dependence structure is the correct general approach, blocking in time can itself induce a form of extrapolation: if the held-out period contains operating conditions or fuel compositions absent from the training period, the model is being asked to predict outside the range it was trained on, and the resulting error estimate conflates genuine generalisation error with extrapolation into an unseen regime. For a seasonally drifting process this is a real possibility, and it means a held-out error must be read with attention to whether the test period was representative of the training period or genuinely novel. This qualification is carried into the interpretation of the validation results in the later chapters.

2.7 Conclusion

This chapter has located the present work within the body of methods it draws on. The target is to be estimated by a soft sensor (Section 2.1), of the physics-anchored kind that combines an empirical estimator with a model of the process, and prior calorific-value estimation has been shown to be either composition-based and offline or, where online, confined to incinerators with added instrumentation — leaving online recovery in a closed-loop cement calciner from

existing signals as the open problem. The empirical estimator will be drawn from the gradient-boosting and linear model classes (Section 2.2), with the choice between them governed by the bias-variance trade-off and by the parsimony considerations that make that choice a matter of regime rather than of principle. The supervision available is of an unusual kind — a single daily aggregate labelling a sequence of higher-rate observations (Section 2.3), for which multiple-instance learning provides the closest precedent, though not a directly applicable method. The process model that anchors the estimator is a digital twin in the broad industrial sense (Section 2.4), corresponding, under a stricter taxonomy, to a digital model calibrated offline. And the question of whether the target is recoverable at all is one of observability under closed-loop control (Section 2.5), a question the identification literature shows, by analogy, to be genuinely non-trivial when feedback can mask the very signal one wishes to estimate.

Two themes run through this review and shape the chapters that follow. The first is that no single off-the-shelf method fits the problem cleanly: each of the established frameworks — soft sensing, gradient boosting, multiple-instance learning, digital twins, multi-rate estimation — supplies part of the apparatus, but the particular combination of a delayed aggregate label, a non-stationary process, and a closed-loop plant means the methods must be adapted rather than applied directly. The second is that the central difficulty is not, in the first instance, one of model choice but of observability: before asking which model estimates the target best, one must ask whether the target is present in the measured signals at all, and under what conditions. It is this question that the next part takes up, beginning with the data and instrumentation available at the Oggaz plant.

Part II

Recovering the Signal

Chapter 3

Data, Instrumentation and Methodology

Part I established what is to be estimated and why, and surveyed the methods available for estimating it. Part II turns to the specific case: the Oggaz plant, the signals it records, and what can be built from them. This chapter is the foundation for everything that follows. It describes the data available, where it comes from, at what rate, and in what form — the instrumentation that produces it, and the methodology by which a daily laboratory value and a stream of online measurements are brought together into a dataset that a model can learn from.

The account is deliberately concrete. The viability of the entire approach rests on the match between two records that were never designed to be used together: a laboratory log of fuel calorific value, kept for quality-control purposes, and a process historian recording the plant's online signals for operational monitoring. The rate at which each is sampled, the period each covers, and the degree to which they overlap determine how much usable supervised data exists — and, as later sections show, that quantity is considerably smaller than the raw label count suggests. Establishing these facts precisely, before any model is built, is what allows the modelling chapters to make honest claims about what was achievable and what was not.

The chapter proceeds as follows. Section 3.1 describes the process and its instrumentation: the signals recorded by the control system, the rates at which they are available, and the distinction between what is measured online and what is measured only in the laboratory. Section 3.2 describes the laboratory measurement of calorific value and the form the labels take. Section 3.3 describes the construction of the supervised dataset from the two records and quantifies the usable overlap. Section 3.4 sets out the validation methodology adopted given the non-stationary character of the process. A final section summarises the data situation that the modelling chapters inherit.

3.1 The process and its instrumentation

The estimation problem is defined by what the plant measures and what it does not. This section describes the signals available from the Oggaz calciner and its associated systems, distinguishing

those recorded continuously by the control system from the single quantity, the fuel calorific value — that is available only through delayed laboratory analysis.

3.1.1 Online process signals

The calciner and its supporting equipment are monitored by a Siemens CEMAT control system, which records the plant's operating signals continuously and archives them in a process historian. The signals relevant to fuel-quality estimation are those that respond, directly or indirectly, to the energy released by the alternative fuel in the calciner: the calciner and preheater temperatures, in particular the temperature at the lowest cyclone stage that the control system regulates most closely; the natural-gas flow to the calciner, which the control system manipulates to maintain that temperature; the flow of the alternative fuel itself; and the associated combustion-air and draught signals. These are read from the control system at the rate at which it logs them, giving a near-continuous record of the calciner's thermal state and of the control actions taken to maintain it.

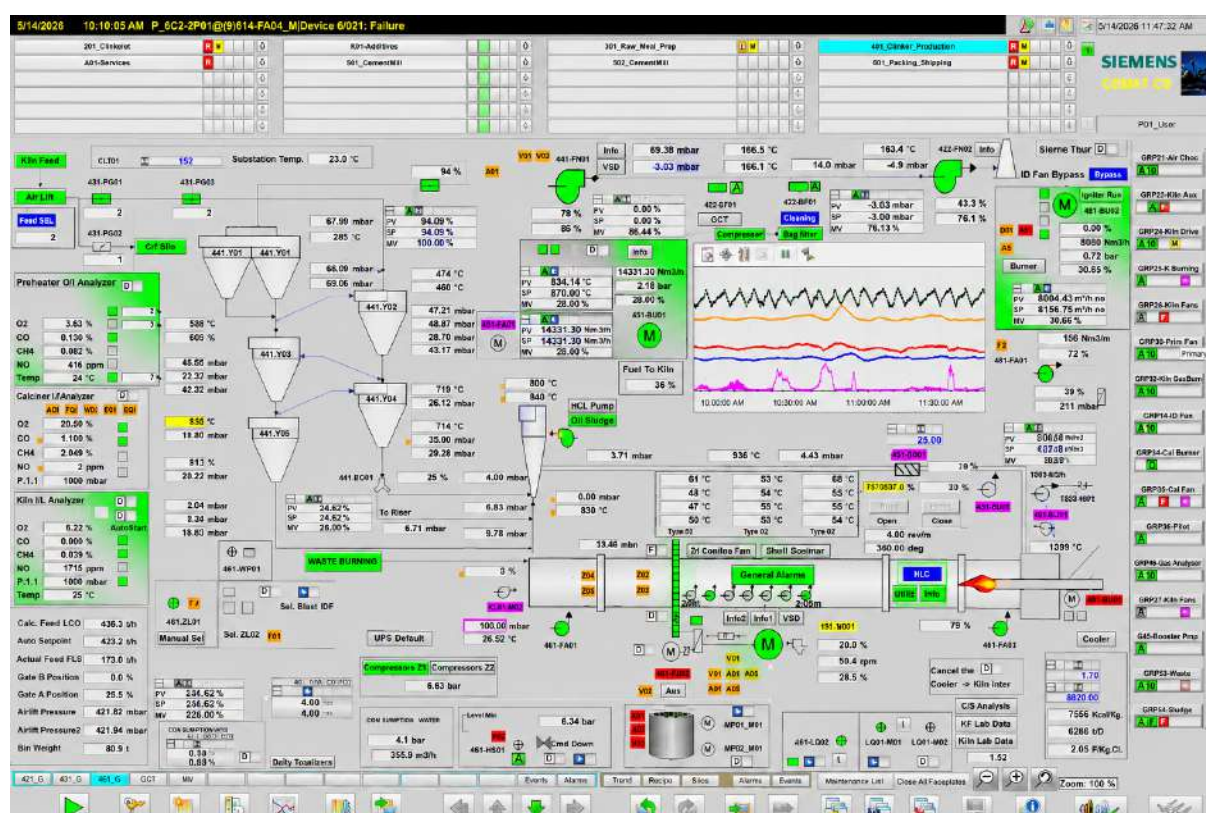


Figure 3.1. The Siemens CEMAT calciner control faceplate.

The defining characteristic of these signals, for the present purpose, is that they are abundant and fast relative to the quantity to be estimated. Where the target is available once per day, the online signals are available continuously, which is what makes an estimation approach conceivable at all: the information used to infer fuel quality must come from these signals, since they are the only continuous window onto the process the fuel is driving.

3.1.2 The alternative-fuel injection system

The alternative fuel is delivered to the calciner by a dedicated injection system at the Sludge Treatment Station, where the blended fuel is stored and pumped to the calciner by a Putzmeister positive-displacement pump. The control system records the pump's operation and the commanded feed, which together describe the delivery of fuel to the calciner over time.



Figure 3.2. The Putzmeister injection-pump control screen at the Sludge Treatment Station.



Figure 3.3. The positive-displacement pump that injects the alternative-fuel stream into the calciner.

The injection system is thus instrumented at the level of pump operation and commanded feed; the question of what these signals establish about the delivered fuel mass is taken up in the later chapter on observability.

3.1.3 What is measured online and what is not

The instrumentation described above shares one feature: it measures the *process*, not the *fuel*. The temperatures, gas flow, and pump signals describe how the calciner is behaving and how it is being controlled, but none of them measures the property of interest — the calorific value of the alternative fuel, directly. That property is established only by removing a sample from the fuel stream and analysing it in the laboratory, at a rate and with a delay described in the next section.

This is the gap the thesis addresses. A quantity that matters to the operation of the plant is measured continuously nowhere and directly nowhere online; it is present in the process signals only through its effect on the calciner, and only to the extent that effect is not masked by the control system that regulates the calciner. Whether that indirect trace is sufficient to estimate the fuel's calorific value is the question the remaining chapters take up.

3.1.4 The recorded signals: available, retained, and used

The control system logs a large number of tags, not all of which bear on fuel-quality estimation. Before any modelling, the raw historian export was reduced to a consistent set of signals on a common one-minute time base. This section records what the export contained, what was kept, and on what grounds, so that the inputs available to the later chapters are fully traceable.

The raw export carried on the order of thirty columns. They fall into a few groups: kiln operation (running hours, kiln and calciner gas flows, kiln feed rate, clinker production, kiln drive speed and torque, ID-fan speed), the calciner and preheater thermal state (calciner temperature, the lowest-cyclone C5 temperature and pressure, burning-zone temperature, preheater-outlet and secondary-air temperatures), combustion and emissions readings (oxygen, carbon monoxide, and nitrogen oxide at the calciner, kiln, and preheater exit, and methane readings), the alternative-fuel feed, and a number of derived or administrative fields (thermal energy, specific thermal-energy consumption, and unlabelled spare columns).

From this set, the signals retained for the work are those that respond, directly or indirectly, to the energy released by the alternative fuel in the calciner, together with the variables needed to control for production rate. These are the calciner gas flow, the C5 temperature and its setpoint, the alternative-fuel feed, the kiln feed rate, and the clinker production, with the calciner oxygen retained as a combustion indicator. The remaining columns were set aside for one of three reasons. Some are not relevant to the calciner energy balance, such as the kiln drive torque and the ID-fan speed. Some are redundant derived quantities computed by the control system from signals already retained, such as the specific thermal-energy consumption. And some are administrative or empty, such as the unlabelled spare columns and the kiln running-hours counter.

Harmonising the export required resolving inconsistencies in the tag names. Several raw headers mapped to the same canonical signal, and a small number of header collisions, where unrelated tags shared a name, were resolved by inspection before the affected columns were assigned. Gaps in the retained signals were treated conservatively: short gaps, of up to fifteen minutes, were filled by holding the last value, while longer gaps were left missing and flagged rather than filled, so that genuine outages were never disguised as data. The result is a clean one-minute record of the retained signals over the historian period, which is the record the audit of Chapter 4 and the modelling that follows are built on.

Table 3.1 summarises the disposition of the principal signals.

Table 3.1. Disposition of the principal recorded signals. Retained signals are those bearing on the calciner energy balance and on production-rate control; the rest are set aside as irrelevant, redundant, or administrative.

Signal	Role	Disposition
Calciner gas flow	Controller output; primary response to fuel energy	Retained
C5 (lowest-cyclone) temperature	Regulated process variable	Retained
C5 temperature setpoint	Control reference	Retained
Alternative-fuel feed	The disturbance input	Retained
Kiln feed rate	Production-rate control	Retained
Clinker production	Production-rate control	Retained
Calciner oxygen	Combustion indicator	Retained
Calciner / burning-zone / preheater temperatures	Secondary thermal state	Retained as available
CO, NO _x , CH ₄ readings	Emissions; not used for estimation	Set aside
Kiln drive speed, torque, ID-fan speed	Mechanical; not in the energy balance	Set aside
Specific thermal-energy consumption	Redundant derived quantity	Set aside
Running hours, spare/unlabelled columns	Administrative or empty	Discarded

3.2 Laboratory measurement of calorific value

The calorific value that the soft sensor estimates is established by laboratory analysis of a physical sample of the fuel. This section describes how that measurement is made and how the sample it is made on is formed, because the character of the label, as much as the character of the input signals, shapes what the estimation problem is.

3.2.1 The measurement

The calorific value of the alternative fuel is measured by bomb calorimetry: a prepared portion of the fuel is combusted in oxygen inside a sealed vessel, and the heat released is determined from the resulting temperature rise. The value is reported on a net basis — the calorific value after accounting for the heat carried away by the water formed in combustion and by the fuel's own moisture — which is the quantity relevant to the energy actually available in the calciner. The measurement itself is standard; what makes it consequential for the estimation problem is not how the calorimeter works but what it is applied to, and how often.

One limitation of this basis must be flagged here, because it returns later as a quantified constraint. The net value accounts for the hydrogen-combustion water, but the as-fired energy available in the calciner also depends on the fuel's own moisture, which absorbs latent heat as it vaporises on injection. Establishing the as-fired value therefore requires a per-sample moisture measurement, and as the data chapter and the audit will show, that measurement is available for only a small minority of days. A systematic gap between the laboratory assay and the in-situ combustion reality is thus present from the outset; its size is quantified in Section 4.4.

3.2.2 The sample: grab sampling and the daily composite

The measurement is applied not to an instantaneous sample but to a daily composite. The fuel reaching the calciner is not uniform in time, so to obtain a single representative figure for a day the plant accumulates grab samples taken from the feed line through the day, mixes them, and analyses a portion of the resulting composite.



Figure 3.4. The alternative-fuel sampling station. At regular intervals a valve on the feed line opens and a grab sample is collected; the day's grabs are combined into a composite, a portion of which is analysed in the laboratory.

This procedure fixes what the resulting number means. The daily calorific value is an *average over the day's fuel*, integrated across whatever variation occurred between grabs, rather than a measurement of the fuel at any single moment. It is this averaging that makes the label an aggregate over the day, the property identified in Section 2.3 as the defining feature of the supervision

available here — and that governs how it can, and cannot, be related to the fast online signals recorded over the same day.

3.3 Constructing the supervised dataset

The laboratory log of calorific value and the process historian of online signals were kept for different purposes and over different periods. A supervised dataset exists only where they overlap: where a daily calorific value can be paired with the online signals recorded over the same day. This section establishes how much such data exists — and it is considerably less than the raw label count suggests, and what kind of label it consists of.

3.3.1 Two records, one overlap

The laboratory record spans a longer period than the process historian. Figure 3.5 places the two on a common timeline. The laboratory labels extend across roughly eighteen months, from late 2024 into mid-2026, but the process historian — the archived online record against which a label must be matched to form a training example — begins only in December 2025. A label dated before the historian record starts has no online signals to be paired with, however valid the measurement itself; it cannot contribute a supervised example.

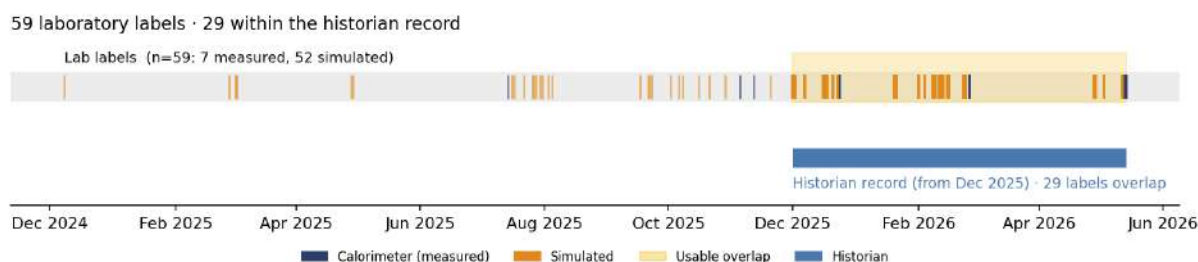


Figure 3.5. The laboratory labels and the process historian on a common timeline. The laboratory labels (top) span late 2024 to mid-2026; the historian record (bottom) begins only in December 2025. Only the labels coinciding with active alternative-fuel feeding and falling within the historian period can be paired with online signals to form supervised examples; a label on a day without active ASF feed carries no disturbance to read and is excluded.

The consequence is visible directly in the figure. Of the laboratory labels collected, 59 are usable; these establish the real calorific-value distribution and the monthly calibration the twin requires to reproduce the seasonal cycle across a full year. The online historian record, which begins in December 2025, overlaps 29 of them, while the remainder predate the online data. Both sets matter for different purposes: all 59 inform the seasonal calibration the twin is built on, while the 29 with paired online signals are those against which real disturbance events could in principle be matched. This is the first and hardest constraint on the work: the quantity of supervised data is fixed not by how many laboratory analyses were performed, but by how many of them coincide with a period of recorded online signals.

3.3.2 The composition of the label set

The 59 usable labels are not all of one kind, and the distinction is the central fact about the ground truth available to this work. Not every day carries a calorimeter result, and the laboratory record fills the gaps with an estimate. Of the 59 daily labels, only **7 are direct calorimeter measurements** of the day's composite; the remaining **52 are produced by the plant's own simulator** — a recipe-forward estimate, generated on days without a calorimeter result, which computes the composite calorific value from the day's blend as the mass-weighted average of its components,

$$CV = \frac{\sum_i m_i CV_i}{\sum_i m_i} \quad (3.1)$$

where m_i is the mass of component i in the day's blend and CV_i its reference calorific value.

Table 3.2. The 59 daily labels by provenance.

Provenance	Meaning	Count
Calorimetric	A direct bomb-calorimeter measurement of the day's composite.	7
Plant-simulated	The plant's own recipe-forward estimate, used on days without a calorimeter result.	52

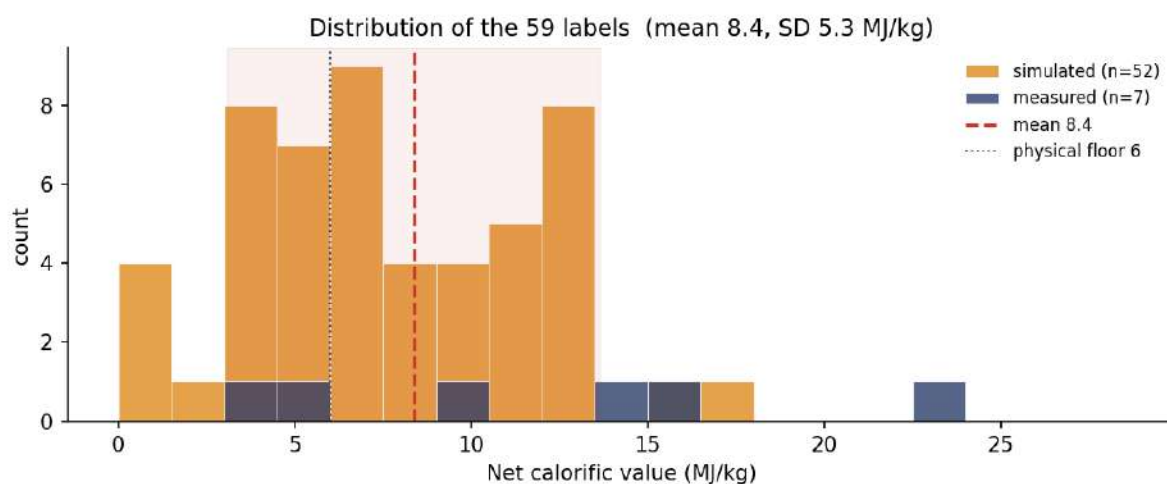


Figure 3.6. Distribution of the 59 labels by net calorific value, coloured by provenance (calorimeter and plant-simulated). The dashed line marks the mean and the shaded band one standard deviation; the dotted line marks the 3 MJ/kg physical floor below which five simulated recipe artifacts fall.

The implication must be stated plainly: the great majority of the labels are not measurements of the burned fuel but estimates of it. A plant-simulated label reflects the assumed recipe and the

reference calorific values built into the plant's simulator, not a physical measurement of what was fed; where it serves as ground truth, the model is in part being trained against, and compared to, the plant's recipe model rather than the fuel itself. This distinction is preserved through the analysis and revisited wherever it bears on the interpretation of results.

One point of terminology must be fixed here to prevent a confusion that would otherwise run through the thesis. The *plant simulator* described above is a pre-existing recipe calculator internal to the plant's laboratory workflow, used here only as a source of labels. It is entirely distinct from the *digital twin* developed later in this work (Chapter 5), which is a process model built to reproduce the calciner's dynamic response. The two are kept terminologically separate throughout: "plant-simulated" always refers to a label's origin, never to the modelling done in this thesis.

3.3.3 The dataset the modelling inherits

The supervised dataset on which the modelling rests is thus 59 irregularly spaced daily labels, each an aggregate over its fuel-day, the great majority of them recipe-forward simulator estimates rather than direct measurements, and reported on a net basis. Excluding five simulated values that fall below the 3 MJ/kg physical floor as recipe artifacts, the 54 plausible labels have a mean of 9.1 MJ/kg and a standard deviation of 5.0, spanning roughly 3.5 to 28.6 MJ/kg; this is the range against which the recovery errors reported later, of order 2.5 MJ/kg, should be read. Moisture is recorded on only five of the 59 fuel-days, which is why the working basis throughout is net rather than as-fired. Figure 3.6 shows the distribution of the labels by provenance. The labels span a wide range of calorific values, reflecting both the genuine heterogeneity of the fuel and the spread introduced by the simulator.

This is a small dataset by the standards of the methods reviewed in Chapter 2, and its size and composition together are the single most important fact about the modelling problem. It is what makes the bias-variance and parsimony considerations of Section 2.2 not an academic point but the governing constraint: with a labelled sample this small — and one whose ground truth is itself largely generated by the plant's recipe estimator rather than measured by calorimetry, the complexity an estimator can support is sharply limited, and the question of how much can be estimated at all, and against what, comes to the fore.

This is the data situation the modelling chapters inherit. What can be recovered from it, and what the limits are, is the subject of the chapters that follow.

3.4 Validation methodology

A soft sensor is only as credible as the procedure used to estimate its error, and with a dataset as small as the one just described, that procedure is not a formality but a decision that shapes what can honestly be claimed. This section sets out how the estimators in the following chapters are validated, and why the standard approach is inappropriate for this process.

3.4.1 Why random cross-validation does not apply

The default procedure for estimating a model's error is k-fold cross-validation: the data are partitioned at random into folds, each fold is held out in turn, and the errors are averaged. This rests on an assumption — that the observations are independent and identically distributed — that does not hold here. The labels are a time series: successive fuel-days are not independent draws from a fixed distribution, because the fuel itself changes over the period of record.

That change is visible directly in the data. Figure 3.7 shows the monthly presence of each component in the alternative-fuel blend over the record. The composition is not stationary: components enter and leave the mix from month to month, and the relative weight of those present shifts as the available waste streams change through the year. The fuel burned in one part of the record is, in its composition, a materially different fuel from that burned in another.

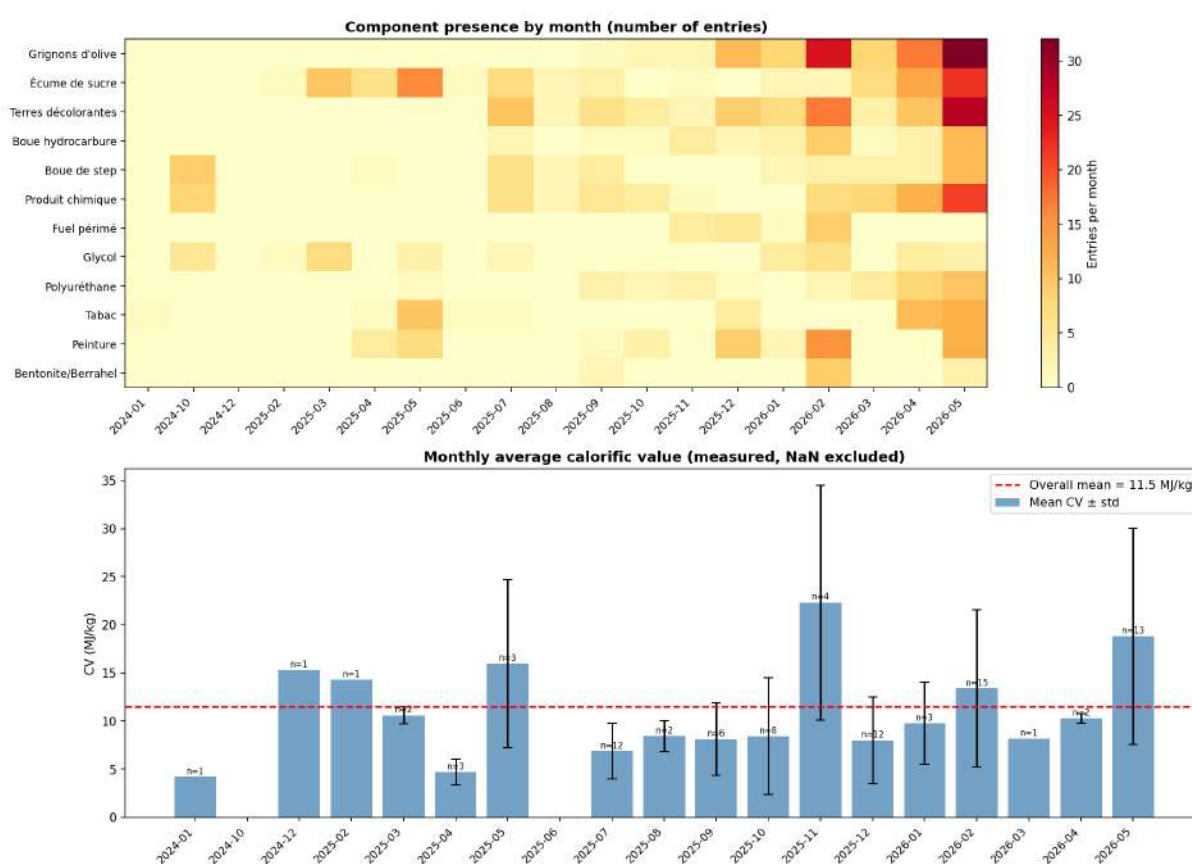


Figure 3.7. Monthly presence of each component in the alternative-fuel blend across the record. The composition shifts from month to month as the available waste streams change, so the fuel is non-stationary over the period.

For a non-stationary series, random cross-validation is not merely suboptimal but unreliable. By drawing training and test points at random from across the whole record, it allows the model to be tested on a period whose fuel composition it has already seen in training, an optimistic situation that does not correspond to how the sensor would actually be used, where it must estimate the calorific value of fuel from a period after the one it was trained on. The literature on time-series evaluation is explicit on this point: Bergmeir and Benítez (2012) [27] show that the independence assumption underlying cross-validation is invalidated by temporal dependence,

and that reserving a contiguous block from the end of the record for testing removes the difficulty by respecting the temporal order.

3.4.2 Out-of-sample evaluation on held-out later periods

The procedure adopted here follows from that. Rather than partition the labels at random, the estimators are trained on an earlier portion of the record and evaluated on a later portion held out entirely from training — an out-of-sample, forward-in-time evaluation that mirrors the intended use of the sensor and respects the temporal order of the data. This choice is directly supported by the empirical literature: Cerqueira, Torgo and Mozetič [28], comparing evaluation procedures across many real-world series, find that for non-stationary series the most accurate error estimates are produced by out-of-sample evaluation on held-out later periods rather than by cross-validation. Since the present process is non-stationary by construction, this is the regime their result addresses.

3.4.3 A caveat the held-out scheme carries

This scheme is the correct one, but it carries a limitation that must be stated rather than hidden, because it bears directly on how the later results are read. Roberts et al. (2017) [29] show that while blocking data to respect temporal dependence is the right general approach, blocking in time can itself induce a form of extrapolation: if the held-out later period contains fuel compositions absent from the training period, the model is being asked to predict outside the range it was trained on, and the resulting error then conflates genuine generalisation error with extrapolation into an unseen regime. For a fuel whose composition drifts as this one does, that is a real possibility. A held-out error must therefore be read with attention to whether the test period was representative of the training period or genuinely novel — a qualification carried into the interpretation of the results in the chapters that follow.

3.5 The 2023 no-ASF reference period

The steady-state analysis of Chapter 4 rests on a comparison against a baseline in which no alternative fuel was fired, so that the gas demand under fossil-only operation can be established and any displacement by the alternative fuel measured against it. That baseline is drawn from a distinct period of record, and because it underpins a central conclusion it is described here rather than assumed later.

The reference period is January 2023. It is taken from the same plant historian as the main record, exported separately, and it was confirmed by the plant engineers to be a month in which no alternative fuel was co-fired; the recorded ASF feed is zero throughout, with no minute showing active feed. It therefore provides exactly the fossil-only reference the main 2025–2026 record cannot supply, since alternative fuel is fired throughout the latter.

The baseline gas figure used throughout is not the raw monthly average but the calciner gas

restricted to matched steady-state production, filtered to the same kiln-feed, gas and oxygen bands and running conditions used elsewhere so that the comparison is like-for-like. Under that restriction the no-ASF calciner gas is 144.2 MW with a standard deviation of 7.3 MW over roughly twenty-four thousand one-minute samples. This standard deviation is the noise floor referred to throughout the thesis: the run-to-run variability of the steady-state gas signal under fossil-only operation, and therefore the level against which any fuel-quality signal must compete.

One question this reference raises is whether a baseline established in early 2023 remains valid for a study period in 2025–2026, given that the plant may have drifted between them. This was tested directly. Using the 2023 baseline to predict the calciner gas on the no-ASF minutes available in the 2026 record gives a mean bias of about +1.2% of the baseline gas, with a prediction error comparable to the baseline’s own run-to-run variability. The baseline therefore transfers across the two-year gap: the fossil-only gas demand at matched production is essentially unchanged, and the 2023 reference is a defensible standard for the later period. This transfer test is what licenses the use of a 2023 baseline in the 2025–2026 analysis, and it forecloses the obvious objection that the two periods are not comparable.

This completes the description of the data and the methodology. The dataset is small, its labels are largely simulator estimates on a net basis, its supervision is an aggregate over each fuel-day, and its evaluation must respect a non-stationary process by holding out later periods. These are the conditions under which the question posed in Part I, whether, and how well, the fuel’s calorific value can be recovered from the available signals — is taken up in Part II’s modelling chapters.

Chapter 4

From Steady-State Observability to the Disturbance Response

The previous chapter established what data exists and in what form. It did not establish whether the quantity of interest — the calorific value of the alternative fuel, can be recovered from that data at all. That is the question this chapter takes up, and it takes it up before any predictive model is built, because the answer determines what kind of model is worth building.

The question is one of observability: whether the fuel-quality signal leaves a recoverable trace in the measured process signals. It is sharpened by a feature of the plant established in Part I — the calciner runs under closed-loop control, with the natural-gas flow adjusted automatically to hold the calciner temperature at its setpoint. A controller that successfully regulates temperature does so by absorbing disturbances, and a disturbance that the controller absorbs is, by that very fact, partly removed from the signals one would use to detect it. Whether enough of the fuel-quality signal survives this regulation to be recovered, and in what part of the process behaviour it survives, is what the audit must determine.

The chapter proceeds as an investigation rather than a demonstration. Section 4.1 examines the most natural place to look for the signal — the steady-state relationship between fuel quality and the gas required to maintain temperature, and finds it weak. Section 4.2 reconsiders the problem in light of that result, asking not whether the signal is present on average but where in the process behaviour it is concentrated. Section 4.3 identifies the disturbance response — the transient excursion that follows a change in the fuel feed — as the part of the record where the signal is recoverable, and characterises it. The chapter closes by stating the constraints this analysis hands forward: what the later modelling can rely on, and what it cannot. Figure 4.1 sets out the causal chain the audit follows and its central conclusion in advance.

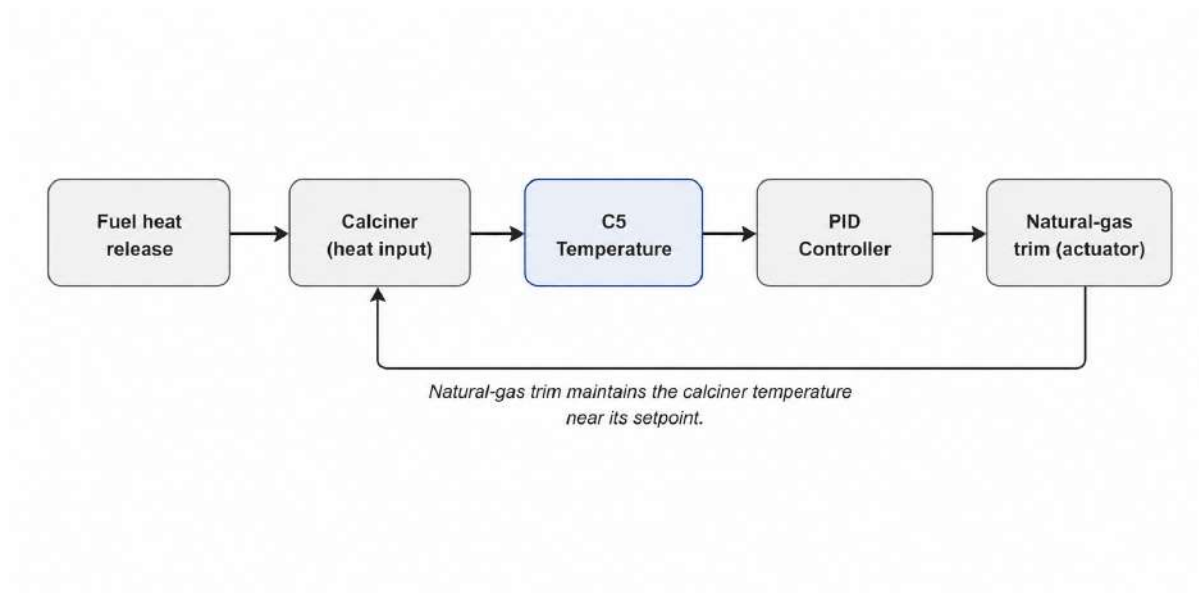


Figure 4.1. Control architecture of the calciner temperature regulation loop.

4.1 The steady-state channel

The most direct way to recover fuel quality from the process is through the energy balance at steady state. When the calciner runs under stable conditions, its heat demand is met by two fuels: the natural gas, which is measured, and the alternative fuel, whose contribution depends on its calorific value. If the alternative fuel carries less energy, the control system must supply more gas to hold the temperature; if it carries more, less gas is needed. The gas required at a matched production rate should therefore fall as fuel quality rises, and if that relationship were strong, the estimation problem would be largely solved at the level of daily averages.

This relationship can be made precise. At steady state, the calciner's heat demand is fixed by the production rate; the natural gas supplies whatever fraction of it the alternative fuel does not. The energy the alternative fuel contributes is then the difference between the gas that would be required with no alternative fuel — a baseline demand, written G_0 — and the gas actually used, G . Dividing that energy difference by the mass of alternative fuel fed, $\dot{m}$, gives the apparent calorific value the fuel must have to account for the gas it displaced:

$$CV_{\text{apparent}}(t) \approx \frac{\Delta \dot{V}_{\text{gas}}(t) CV_{\text{gas}}}{\dot{m}_{\text{ASF}}(t)} \quad (4.1)$$

where $\Delta \dot{V}_{\text{gas}}$ is the displacement in natural-gas volumetric flow relative to the no-ASF baseline, the quantity the control system actually measures, in Nm^3/h ; CV_{gas} is the calorific value of natural gas, taken as the constant $36 \text{ MJ}/\text{Nm}^3$; and $\dot{m}_{\text{ASF}}$ is the alternative-fuel mass rate. Writing the numerator this way keeps explicit what is measured and what is assumed: the plant reads a gas volume, and converting it to energy rests on the assumed constant CV_{gas} .

This is the steady-state estimator in its simplest form, and it is the relationship the rest of this

section tests. It requires three things to be usable: a reliable baseline G_0 , a gas displacement $G_0 - G$ large enough to measure against the variability of the gas signal, and a known fuel mass $\dot{m}$. The analysis below examines each, and finds the displacement term, the numerator — too small to resolve at steady state. The mass term in the denominator, known only to within a density-dependent bound, is quantified later as a constraint in its own right (Section 4.4). The test requires first resolving an artifact in the feed signal that would otherwise corrupt any comparison.

4.1.1 A reporting-scale artifact in the feed signal

The recorded alternative-fuel feed does not sit on a single scale across the record. Figure 4.2 plots the daily median of the non-zero feed on a logarithmic axis: through the end of February the feed is reported around 130 in its raw units, and on 1 March it drops by more than an order of magnitude to around 7, where it remains for the rest of the record. The plant's operation did not change discontinuously on that date; the *reporting* did.

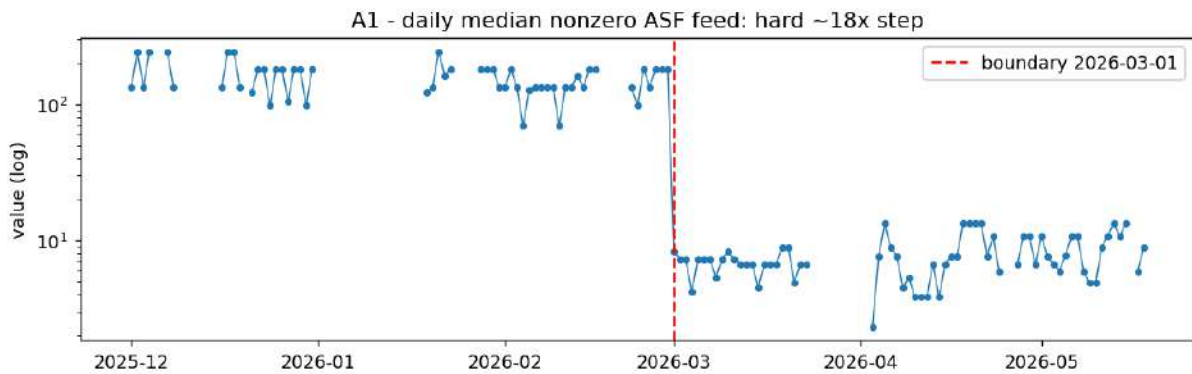


Figure 4.2. Daily median of the non-zero alternative-fuel feed, logarithmic axis. The signal sits roughly eighteen times higher before 1 March 2026 than after, a hard reporting-scale step rather than a change in plant operation.

That this is a reporting artifact and not a real change is established by a physical test. Comparing the two periods at matched production — the same kiln-feed band, the same clinker rate, the calciner gas and the clinker output are statistically identical across the 1 March boundary, while only the feed value jumps. A genuine eighteen-fold increase in fuel delivery would have to show in the gas demand and the production; it does not. The factor itself is obtained empirically, as the ratio of the median recorded feed in the two regimes at matched production: 132.6 t/h before the boundary against 7.3 t/h after, a ratio of 18.2. Dividing the pre-March feed by this factor brings the two periods onto a common scale, and Figure 4.3 shows the gas-versus-feed relationship and the feed distributions of the two regimes coinciding after the correction. Because the gas demand and clinker output are statistically identical across the boundary while only the recorded feed jumps, the factor reflects a change of recording unit rather than of process, and it is applied to the pre-March data throughout.

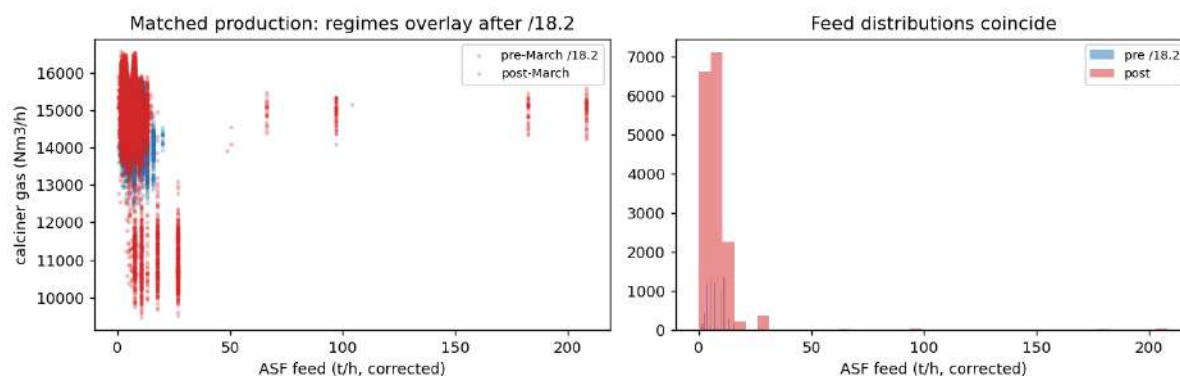


Figure 4.3. After dividing the pre-March feed by 18.2, the two reporting regimes overlay — both the gas-versus-feed scatter (left) and the feed distributions (right) coincide, confirming a single underlying process recorded on two scales.

With the feed on one scale, the steady-state test can be made.

4.1.2 The steady-state displacement is weak

If the alternative fuel measurably displaces gas, the calciner gas at matched production should sit below the no-ASF baseline by an amount that grows with the fuel’s energy content. The 2023 record, in which no alternative fuel was fired, provides that baseline: a mean calciner gas of 144.2 MW, established and validated for transfer to the study period in Section 3.5. Figure 4.4 compares it with the two alternative-fuel periods at matched production, and Table 4.1 gives the underlying figures.

Table 4.1. Calciner gas at matched steady-state production: the no-ASF 2023 reference against the two alternative-fuel periods.

Period	Mean gas (MW)	Std (MW)	n (minutes)
2023 no-ASF reference	144.2	7.3	23976
High-ASF (Dec–Feb)	143.6	6.6	12275
Low-ASF (Mar–May)	147.6	10.3	22609

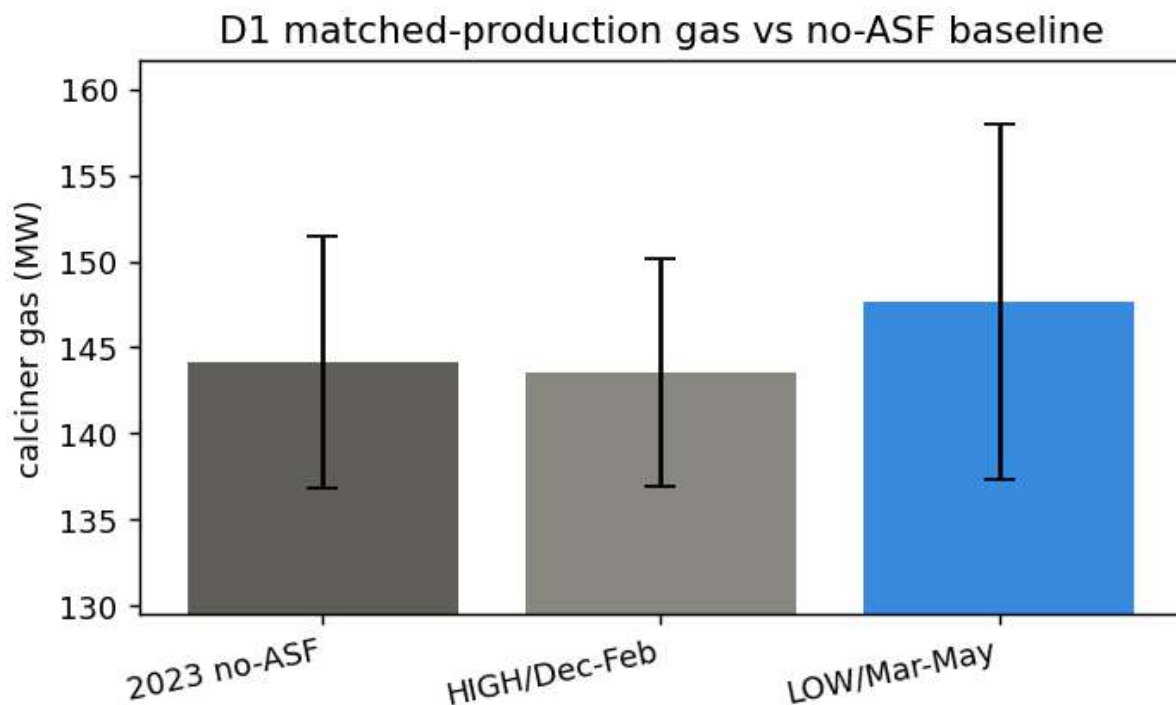


Figure 4.4. Calciner gas at matched production for the no-ASF 2023 reference and the two alternative-fuel periods. The three means differ by less than the run-to-run variability of the gas: the alternative fuel’s displacement of gas is smaller than the noise floor, so the steady-state gas level cannot resolve it.

The displacement is not visible above the noise. The matched-production gas in the alternative-fuel periods differs from the no-ASF baseline by only a few megawatts, while the run-to-run variability of the baseline gas, about 7.3 MW, is itself larger than that difference. This 7.3 MW figure, used throughout the thesis as the noise floor, is defined precisely as the standard deviation of calciner gas at matched steady-state production over the no-ASF January 2023 reference period described in Section 3.5, computed on roughly twenty-four thousand one-minute samples; it is the same reference that fixes the 144.2 MW baseline. Whatever energy the alternative fuel contributes, its signature on the absolute gas demand sits beneath the variability of the gas signal itself: the absolute displacement cannot be read off the steady-state gas level.

A weaker version of the relationship does survive when the comparison is made directly against laboratory calorific value rather than against the baseline. Restricting to the days where a clean fuel-quality label can be paired with the steady-state signals, the daily gas does fall with rising calorific value, Figure 4.5, where the fitted correlation is $r = -0.53$. But the relationship is fragile in a way the figure makes plain: it leans heavily on the single high-calorific-value batch at the right of the plot, near 38 MJ/kg. On the broader clean overlap of process and labels — a larger day-set not dominated by that one point, the correlation falls to -0.287 , and the partial correlation controlling for feed rate to -0.299 , on labels whose calorific-value spread is itself only about 3 MJ/kg.

The trend is therefore real but weak, and rests disproportionately on a single extreme batch. This is the central finding of the steady-state analysis. The signal that the energy balance predicts is present — gas does fall as quality rises — but at the daily, steady-state scale it is weak relative

to the variability of the signals it must be read from, and far too weak to support recovery of the absolute calorific value. The displacement the alternative fuel makes on the average gas demand is, in this process, near the floor of what the gas signal can resolve.

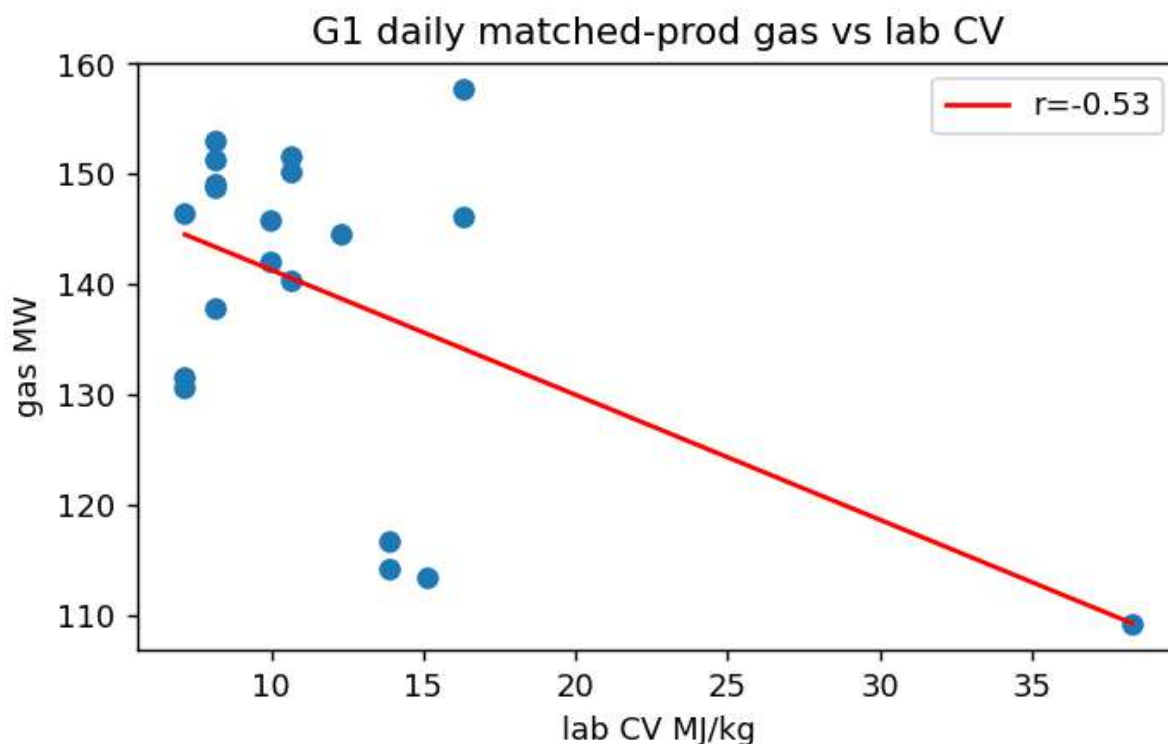


Figure 4.5. Daily matched-production gas against laboratory calorific value. A downward trend is present ($r = -0.53$ on this set) but loose, and it leans heavily on the single high-calorific-value point at the right.

This is not a failure of method but a finding about the process: at steady state, the fuel-quality signal is largely masked. The natural question is then whether it is masked *everywhere*, or whether it is concentrated in some other part of the process behaviour. That is the question the next section takes up.

4.2 Where the signal is not masked

The steady-state result is not that the alternative fuel contributes no energy — it plainly does, but that its contribution does not surface in the average gas demand. This is what the control configuration of the plant would lead one to expect. As established in Section 2.5, the calciner runs under closed-loop control: the natural-gas flow is adjusted continuously to hold the calciner temperature at its setpoint. A controller that regulates temperature well does so precisely by cancelling disturbances — and a change in fuel quality is, to the controller, a disturbance. At steady state the loop has already done its work: it has found whatever gas flow holds the temperature, and the fuel-quality information has been absorbed into that adjustment rather than left visible as a temperature deviation. The better the loop regulates, the more completely the signal is hidden in the settled state.

But a loop cannot cancel a disturbance instantaneously. When the fuel quality changes abruptly, there is an interval — before the controller has caught up, in which the disturbance is visible: the temperature moves, and the gas ramps to bring it back. The information the steady state conceals is exposed in the transient. This reframes where to look. Rather than the average gas demand, the place to seek the fuel-quality signal is the process's *response to a disturbance* — and the cleanest disturbance available in this record is a deliberate one: the moments when the alternative-fuel feed is cut and the calciner must be held on gas alone. The next section examines those events.

4.3 The disturbance response

4.3.1 A single event: the feed cut of 13 May

On the morning of 13 May 2026 the alternative-fuel feed, which had been running at around 18 t/h, was cut to zero and held there. Figure 4.6 shows the three signals over that morning: the feed (top), the calciner gas (middle), and the C5 calciner temperature against its setpoint (bottom).

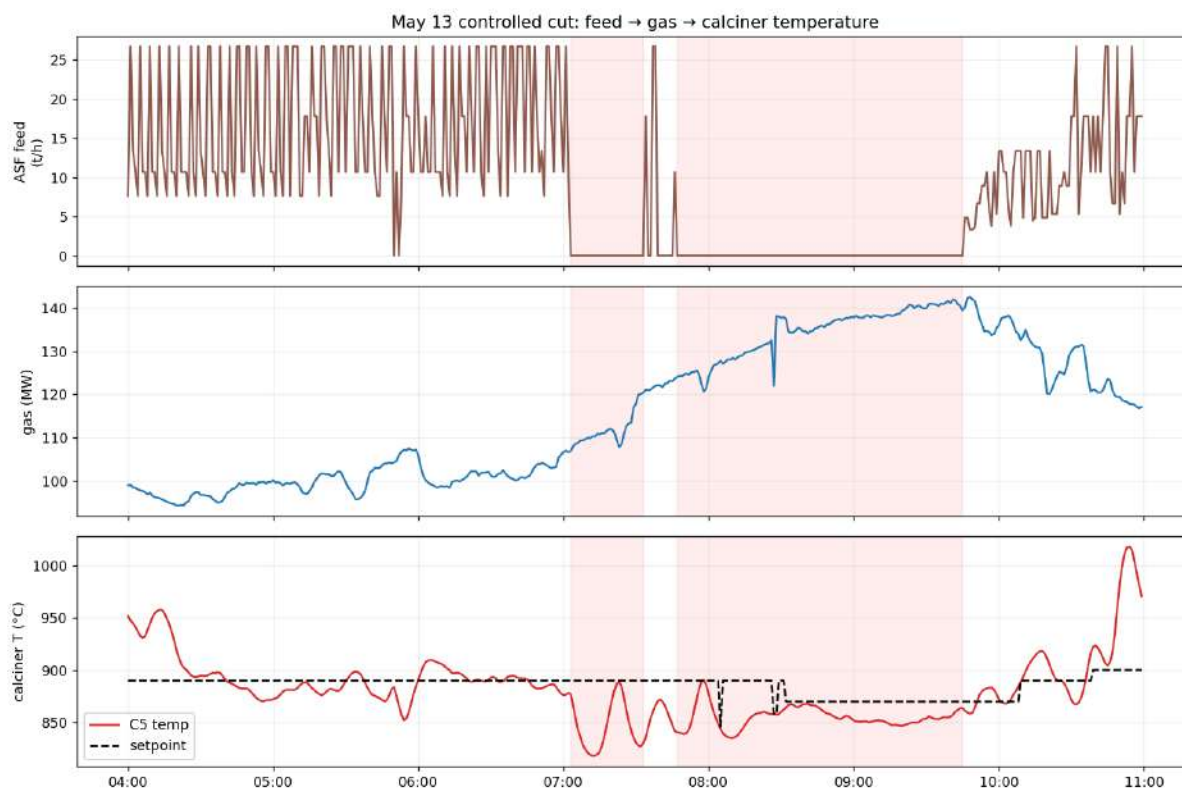


Figure 4.6. The feed cut of 13 May 2026. When the alternative-fuel feed (top) is cut to zero, the calciner temperature (bottom) falls below its setpoint and the control system ramps the gas (middle) to recover it. Shaded bands mark the sustained ASF-off intervals.

The sequence is unambiguous. While the feed is on, the temperature holds near its setpoint and the gas sits around 99 MW. When the feed is cut, the temperature falls — the C5 reading troughs around 818 °C, some 70 °C below setpoint at its worst, and in response the gas ramps upward,

climbing from roughly 99 MW to a peak near 138 MW as the controller works to restore the temperature. When the feed returns, the gas relaxes back.

Read against the steady-state result, this is the same energy balance made visible. The energy the alternative fuel had been supplying did not vanish at steady state; it was simply matched by a gas adjustment too gradual to separate from the noise. Here, forced into the open by an abrupt cut, the same quantity appears as a large, clearly resolved excursion: the depth of the temperature trough and the size of the gas ramp are a direct measure of how much energy the withdrawn fuel had been contributing. The disturbance response carries the fuel-quality signal that the steady state masks.

4.3.2 Does it repeat? The event inventory

A single event, however clean, proves only that the signal can appear, not that it is a reliable feature of the process. If the disturbance response is to be useful, the same cause must produce the same kind of response across many events. To test this, every sustained feed cut in the record was identified — intervals of at least thirty minutes, preceded by steady matched production and a real feed. Thirteen such events were found; the twelve resolved here are listed in Table 4.2, ordered by the depth of the temperature excursion, and Figure 4.7 plots that depth against the feed level immediately before each cut.

Table 4.2. Sustained feed-cut events, ordered by temperature-excursion depth. The two 13 May entries are distinct cut intervals on the same day, a short morning cut and a long afternoon cut. Only three events carry a fuel-CV label; the rest occurred on days without laboratory analysis.

Date (start)	Duration (min)	Feed before (t/h)	Gas swing (MW)	Temp drop (°C)	Fuel-CV label
15 May	78	16.1	31.6	−83.5	—
13 May (07:03)	30	18.3	18.4	−71.9	15.0
14 May	112	16.1	43.8	−63.6	14.2
14 Apr	327	5.4	−47.7	−47.4	—
13 May (13:49)	1234	18.3	51.9	−46.9	15.0
11 May	1962	7.1	2.1	−29.5	—
21 Mar	784	5.9	6.7	−29.1	—
15 Mar	991	5.4	9.1	−23.9	—
7 Mar	856	5.8	1.9	−20.8	—
10 Mar	48	5.5	5.5	−13.6	—
19 Mar	256	6.8	2.3	−12.5	—
30 Apr	222	8.5	−2.8	−8.9	—

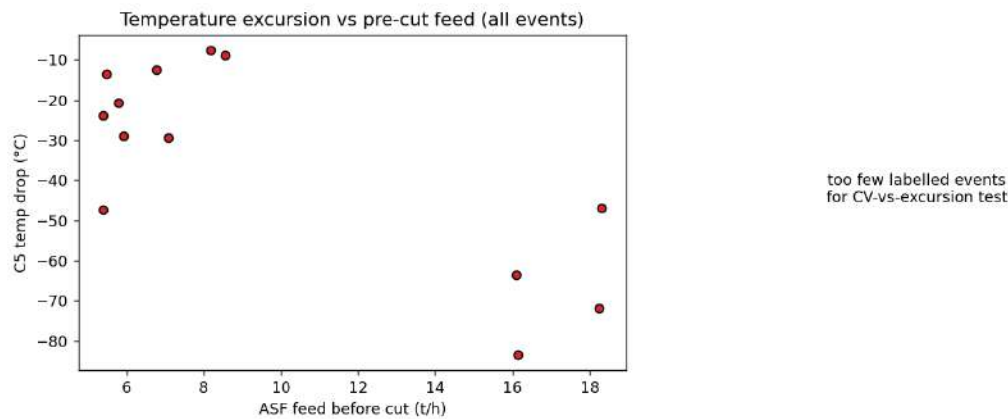


Figure 4.7. Temperature-excursion depth against pre-cut feed level for the feed-cut events. Deeper excursions tend to follow higher pre-cut feed, as a fuel-energy interpretation predicts, though the relationship is loose and the labelled events are few.

Three things are visible in the table and figure together. First, the response repeats: every sustained cut produces a temperature excursion, and the deeper excursions tend to follow the higher pre-cut feed levels, consistent with the reading that the excursion reflects the energy withdrawn. The 13 May morning cut is the clearest single event, but it is not unique; the 14 and 15 May cuts reach comparable or greater depth. The disturbance response is a recurring feature of the process, not a one-off.

Second, the relationship is loose, and the table shows where it breaks down. Two low-feed events, 14 April and 30 April, have a *negative* gas swing: the gas fell rather than rose when the feed was cut, the opposite of the energy-balance reading. The two are not the same kind of event, however, and it would be wrong to treat them alike. The 30 April swing of -2.8 MW is small, well within the 7.3 MW noise floor, and is exactly the kind of inversion expected when a small cut is read against ordinary process variation. The 14 April swing of -47.7 MW is a different matter: at nearly seven times the noise floor it cannot be dismissed as noise, and the energy balance offers no reason for a cut in feed to be accompanied by so large a *fall* in gas. This event is therefore treated as a genuine anomaly rather than as ordinary variation: its cause is not established from the available data, and rather than force an explanation it is flagged as unexplained and excluded from the analysis, along with the low-feed regime it belongs to. What the table does show cleanly is that the interpretable response is a feature of the *large* cuts; at low feed it is either lost in the normal movement of the process, as on 30 April, or, as on 14 April, overtaken by something the record does not explain. This vulnerability cuts in a specific direction. If ordinary process variation can invert the thermodynamic signature at low feed, the same variation is present at high feed, where it does not invert the signature but may still distort its magnitude. The clearest event, 13 May, cannot be assumed free of this: its excursion may combine the feed cut with coincident process movement, and with only three labelled events there is no way to average that contamination out. This is a direct reason passive historical data is insufficient to train an estimator here. It motivates the digital twin of the next chapter, which generates disturbance events under known fuel conditions against a calibrated stochastic noise floor set to the plant's

measured 7.3 MW gas variability, so that a model learns to read the transient signature against the same interference the real plant carries rather than an unrealistically clean background.

Third, and decisively for what can be built on it: of these events, only three carry a fuel-quality label — the two 13 May cuts and the 14 May cut. The remaining events occurred on days with no laboratory analysis of the fuel, so although the *response* is recorded, the fuel quality that produced it is unknown. The disturbance response is therefore clearly observable, but the data needed to *calibrate* it — to turn an excursion into a calorific value, is almost absent. Three labelled events cannot fit and validate a mapping from disturbance response to fuel quality; they can only show that such a mapping is plausible.

4.4 Constraints handed to the digital twin

The audit has located the fuel-quality signal in the disturbance response and shown that the real record holds too few labelled events to exploit it. The next chapter answers this by building a model of the process to generate the missing events. For that model to be credible, the synthetic events it produces must resemble the real ones — which means the model must reproduce the process as the audit found it. This section sets out the quantified constraints the audit hands forward: the mass uncertainty on the feed, the moisture basis of the labels, and the dynamics of the disturbance itself. These are not new analysis; they are the calibration targets the twin must hit.

4.4.1 Feed mass: quantised levels and a density bound

The alternative-fuel feed does not vary continuously. Restricted to its trustworthy active range (0.5–28.6 t/h), it holds at a small number of recurrent levels, listed in Table 4.3. Seven levels account for the great majority of the active feed time, with the two lowest — 4.7 and 7.7 t/h, covering roughly two-thirds of it. The feed is, in effect, a staircase among these levels rather than a smoothly varying signal.

Table 4.3. The seven recurrent levels at which the alternative-fuel feed holds, with the fraction of active time spent at each.

Feed level (t/h)	Share of held time
4.7	38.6%
7.7	27.0%
10.0	5.6%
10.7	7.6%
11.5	5.6%
13.4	12.2%
16.0	1.8%

The feed is recorded as a volumetric rate, but the quantity that matters for the energy balance is mass. Converting one to the other requires the density of the sludge, which is not measured per batch and is known only to lie in the range 1.2–1.5 t/m³. This bound is not a statistical estimate but a derivation from the measurement chain: the recorded feed is a volumetric rate scaled by an assumed density, drawn from a basin of about 260 m³ that empties over roughly twelve hours. Taking the density at the midpoint of its physical range, the half-width of that range, (1.5 – 1.2)/2 divided by the 1.35 midpoint, gives a **±11% uncertainty on the delivered mass**, and because the apparent calorific value is energy divided by mass, this uncertainty maps one-to-one onto the apparent calorific value: a feed whose true mass is 11% higher than assumed yields an apparent calorific value 11% lower for the same energy. The twin must therefore not treat the feed mass as exact; it samples a mass multiplier across the range 0.89–1.11, so that the synthetic data carries the same mass ambiguity the real data does.

4.4.2 Moisture: the labels are on a net basis

The second constraint concerns the basis on which calorific value is expressed. The energy actually available in the calciner is the as-fired value, which requires the fuel’s moisture content to apply the latent-heat penalty for vaporising that water. But moisture is recorded on only a small minority of fuel-days, as Table 4.4 shows.

Table 4.4. Moisture coverage across the fuel-days, and the as-fired penalty where moisture is known.

	Value
Fuel-days with a moisture value	5 of 59 (8%)
Median moisture where present	25.3%
Moisture range	6.3 – 35.0%
Median as-fired penalty	–4.2 MJ/kg ($\approx$ 30% of net CV)

The consequence is significant. Where moisture is known, converting to an as-fired basis lowers the calorific value by a median of 4.2 MJ/kg — about a third of its net value — so moisture is not a minor correction. But with coverage at only 8%, the as-fired value cannot be computed for the real data as a whole, and the working labels are therefore on a net basis throughout, as established in Chapter 3. The twin handles this by construction: because it generates the fuel itself, it knows each batch’s moisture exactly, samples it from the observed distribution (median 25%, range 6–35%), and applies the as-fired penalty internally. It can thus produce labels on either basis, and the gap between them, invisible in the real data — becomes measurable in the synthetic data.

4.4.3 Disturbance dynamics: the shape the twin must reproduce

The first two constraints concern the fuel; the third concerns the process response, and it is what makes the disturbance events realistic rather than instantaneous. Two dynamic limits, measured

from the record, define the shape of a feed-cut event.

The first is how the feed itself moves. It does not glide between levels but holds and steps: the median dwell at a level is about 2 minutes, and 90% of holds last no more than 4 minutes. A synthetic feed must reproduce this staircase, holding briefly and then stepping, rather than varying smoothly — otherwise the disturbances it drives would not resemble the real ones.

The second is how fast the gas can respond. The calciner gas cannot change instantaneously; it ramps at a limited rate, with a median slew of 0.34 MW per minute and a 95th-percentile slew of about 1.56 MW per minute. This rate limit is what gives the disturbance response its shape: when the feed is cut, the gas cannot leap to its new level but climbs over minutes, exactly the ramp seen in Figure 4.6. The twin's controller must be slew-limited near this rate, because that ramp *is* the signal the later model will read; a gas that recovered instantly would leave no excursion to learn from.

Together these four constraints — the quantised feed and its $\pm 11\%$ mass uncertainty, the net-basis labels and the moisture distribution behind them, the feed staircase, and the slew-limited gas response — are the specification the twin is built against. They are what make its synthetic disturbance events resemble the real ones closely enough that a model trained on them can be expected to transfer.

4.5 Conclusion

This chapter set out to determine whether the alternative fuel's calorific value can be recovered from the available signals, and where, and to do so before any estimator was built, because the answer governs what is worth building.

The steady-state channel, the natural first candidate, was tested and found weak. Once a reporting-scale artifact in the feed signal was resolved by the proven factor of 18.2, the matched-production gas in the alternative-fuel periods was shown to differ from the no-ASF baseline of 144.2 MW by less than the run-to-run variability of the gas itself. A relationship between daily gas and laboratory calorific value does exist, but it is loose and leans on a single extreme batch. At the daily, steady-state scale the fuel-quality signal is largely masked — as the plant's closed-loop control predicts, since the controller absorbs the disturbance into a gas adjustment too gradual to separate from noise.

The signal is exposed in the transient. When the feed is cut abruptly, the calciner temperature falls and the gas ramps to recover it, and the size of that excursion reflects the energy the withdrawn fuel had been supplying. This disturbance response was shown to be a clear and repeating feature of the process across the recorded feed-cut events — clean for the large cuts, lost in the noise for the small ones. It is clearly observable. What it is not, yet, is calibrated: only three of the events carry a fuel-quality label, far too few to learn a mapping from excursion to calorific value.

This is the audit's central result, and it defines the problem the remaining chapters address. The

fuel-quality signal lives in the disturbance response, but the real record cannot supply enough labelled disturbance events to exploit it. The response is the subject of the next chapter: a calibrated model of the process, a digital twin — built to generate disturbance events under known fuel conditions, and so to provide the labelled examples the plant itself does not. The constraints it must respect have been quantified here: the quantised feed and its $\pm 11\%$ mass uncertainty, the net-basis labels and the moisture distribution behind them, the feed staircase, and the slew-limited gas response. Against that specification, the twin is built.

Chapter 5

Digital Twin of the Calciner

The audit reached a specific impasse. The fuel-quality signal is observable in the disturbance response, but the plant produced only three labelled feed-cut events, all at high calorific value, far too few to learn the mapping from an excursion to a calorific value. This chapter answers that impasse by constructing a model of the calciner that can generate the labelled disturbance events the plant did not: many events, spanning the full range of fuel quality, each with its calorific value known exactly because the model itself sets it. This chapter describes what that model is, how it is built, how it is used to generate a labelled dataset, the constraints from the audit it must satisfy, and the checks it must pass before any estimator is trained on its output.

5.1 What the twin is, and what it is not

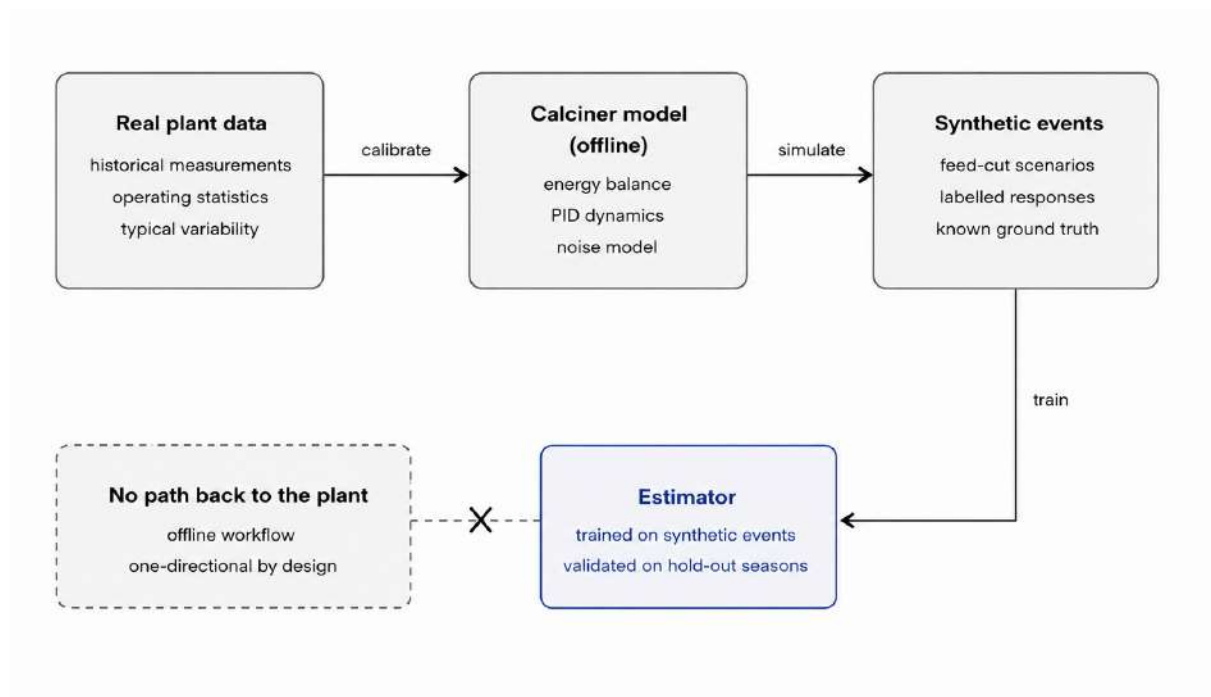


Figure 5.1. The digital twin as a one-directional calibrated forward model. Real plant data calibrate a forward model of the calciner, which generates labelled synthetic disturbance events for training the estimator. No path runs back to the plant: under the Kritzingier taxonomy this offline, one-way artifact is a digital model rather than a bidirectional digital twin.

Section 2.4 distinguished three levels of digital representation after Kritzingier et al. (2018) [20]: a digital model with no automated data exchange, a digital shadow with one-way flow from plant to model, and a digital twin with automated flow in both directions. The choice to calibrate the model offline, rather than couple it bidirectionally to the running plant, is a deliberate engineering decision rather than a limitation worked around. A digital representation that fed signals back to a live calciner holding raw meal above 1400 degrees Celsius would introduce an automated actor into a safety-critical control loop, and generating the thousands of labelled feed-cut events a data-hungry estimator needs would require perturbing clinker production far beyond what any operating plant could permit. The forward model was therefore scoped to do exactly what the estimation problem requires and no more: reproduce, offline, the disturbance dynamics from which calorific value must be recovered. By that taxonomy the artifact built here is a digital model. It is calibrated offline against recorded plant behaviour and then run forward to generate data; it does not receive live data from the plant, nor does it feed estimates back to the control system. The word twin is retained, as set out in Section 2.4, in its broad industrial sense, while its true level of integration is stated plainly here, so that no stronger claim is read into it than the work supports.

What the model reproduces is the calciner’s energy-balance response to fuel. Given a fuel calorific value and a feed rate over time, both set by the modeller and therefore known exactly, it produces the sensor traces a real calciner would show: the calciner gas, the C5 temperature, the oxygen. The unmeasured influences on the process, including raw-meal quality, draught,

ambient conditions, and combustion variability, are not modelled as named mechanisms. They are represented together as a single calibrated stochastic disturbance. This is a deliberate limitation with a sound reason: those variables are not measured at the plant, so no model trained for deployment there could use them as inputs, and modelling them individually would add unfalsifiable detail rather than fidelity. The twin models what the plant can see, and bundles the rest into a noise term whose size is fixed by the data.

The model is therefore a calibrated forward simulator, not a first-principles one. It does not solve the chemistry of clinker formation or the kinetics of combustion. It reproduces, at the level of the measured signals, how the calciner responds when fuel of a given quality is fed and when that feed is disturbed, which is precisely and only what the estimation problem requires.

5.2 The forward model

5.2.1 Method and implementation

The twin is a discrete-time dynamic simulation of the calciner's thermal balance, integrated minute by minute over the period simulated. It was implemented directly in Python using NumPy for the numerical work, without a dedicated discrete-event or process-simulation package: the model is a system of two coupled update equations advanced in a time loop, which NumPy handles efficiently and transparently. This choice keeps the model fully inspectable, with every term in the energy balance and every controller action visible in the code rather than hidden inside a library, which matters for a model whose credibility rests on its being calibrated to specific plant numbers.

The state advanced at each one-minute step is the calciner temperature. Its change is governed by an energy balance: the rate of temperature change is the net heat rate, the heat supplied by gas and alternative fuel minus the process heat demand, divided by a thermal capacitance that sets how fast the temperature responds. In the notation of the model,

$$\frac{dT}{dt} = \frac{(\dot{Q}_{\text{gas}} + \dot{Q}_{\text{ASF}}) - \dot{Q}_{\text{demand}}}{C_{\text{th}}} \quad (5.1)$$

where T is the calciner temperature, $\dot{Q}_{\text{gas}}$ the heat supplied by natural gas, $\dot{Q}_{\text{ASF}}$ the heat supplied by the alternative fuel, $\dot{Q}_{\text{demand}}$ the process heat demand, and C_{th} the thermal capacitance. The alternative-fuel heat $\dot{Q}_{\text{ASF}}$ is the product of its mass rate and calorific value, reduced by the two efficiency terms described below; the gas heat $\dot{Q}_{\text{gas}}$ is set by the controller; and the demand is the baseline heat load plus a temperature-dependent loss term and the stochastic disturbance.

5.2.2 The nonlinearities

A purely linear energy balance would make the calorific value trivially recoverable, which the audit showed it is not. Three nonlinearities, each with a physical basis, make the relationship

between fuel quality and the measured signals realistically indirect.

The first is combustion efficiency as a function of temperature: as the calciner cools, the alternative fuel burns less completely, so a falling temperature reduces the very heat that would arrest the fall. This is represented by an efficiency that drops smoothly as temperature falls below the operating range. The second is moisture: wetter fuel burns less completely, so the burnout efficiency falls with moisture content, reducing the heat delivered per unit of fuel mass. The third is heat loss: the loss from the process rises faster than linearly with temperature, and is defined to be zero at setpoint so that it acts only as a restoring penalty when the temperature departs from its operating point. None of these is a fitted curve dressed as physics; each is a monotone relationship of the expected sign, included so that the mapping from calorific value to excursion is curved and confounded rather than a clean proportionality.

5.2.3 The controller

The gas is set by a controller that mirrors the plant's: a feedforward term that lowers the gas demand by the heat the alternative fuel is expected to supply, plus a proportional-integral feedback term that corrects the temperature error. Crucially, the gas the controller commands cannot change instantaneously. It is slew-limited: the minute-to-minute change in gas is capped at an empirical constraint measured from the real plant, 1.56 MW per minute at the 95th percentile. Whether this limit arises from physical valve-actuation constraints or from the tuning of the plant's own feedback controller is not observable from the available data; what matters here is that enforcing this empirical slew rate in the simulation is what gives the disturbance response its shape. When the feed is cut, the controller cannot leap the gas to its new level; it ramps over minutes, and that slew-limited ramp is the excursion the audit identified as the signal. A controller without this limit would recover the temperature instantly and leave nothing to learn from.

5.3 Constraints from the audit

The twin is built to the specification the audit handed forward in Section 4.4. Each constraint enters the model as a calibration target rather than a free choice.

The **feed** is generated as a quantised, held signal rather than a continuous one. It snaps to the discrete pump levels found in the audit, holds each for a sampled dwell of two to six minutes, steps to another level, and occasionally cuts to zero to create a disturbance event. The level set used is [4.7, 7.7, 10.0, 10.7, 11.5, 13.4, 16.0] t/h, weighted by the dwell-frequencies observed at the plant so that the low levels dominate as they do in the record; the 26.7 level represents the brief high-feed spikes the real feed shows. Figure 5.2 shows a generated day: discrete levels, short dwells, and cut events, with no continuous variation the plant never exhibits.

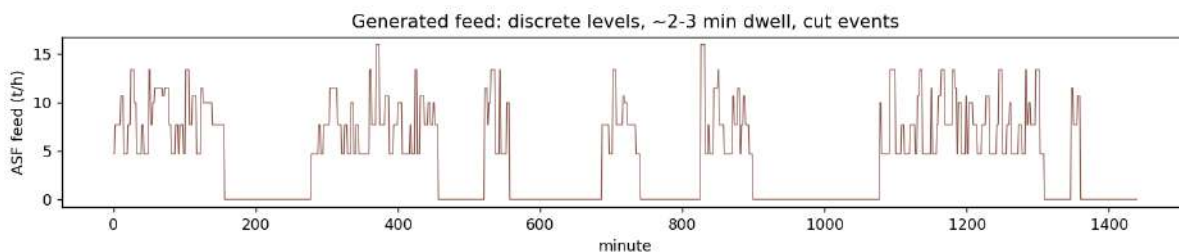


Figure 5.2. A generated feed day. The feed snaps to discrete levels, holds each for two to three minutes, steps, and cuts to zero at intervals to create disturbance events, reproducing the quantised, held character of the real feed.

The **mass uncertainty** enters as a multiplicative factor. Because the delivered mass is known only to within the density-driven $\pm 11\%$ bound established in Section 4.4, each simulated day's feed is multiplied by a random factor drawn from the range 0.89 to 1.11 before it is recorded as the measured feed, so the synthetic data carries the same mass ambiguity as the real data: the model knows the true mass, but the recorded feed does not.

The **moisture** is sampled per day from the observed distribution, centred near the median of 25 percent, and applied through the burnout-efficiency term. Because the twin sets the moisture itself, it can produce the calorific value on either a net or an as-fired basis and knows the gap between them exactly, a gap that on the real data is unmeasurable because moisture is recorded on so few days.

The **disturbance** is generated as a slow random signal, smoothed over an hour so that it represents drifting process conditions rather than minute-to-minute jitter, and scaled so that the unexplained variation in gas at matched feed reproduces the plant's measured noise floor of 7.3 MW. This is the term that makes the recovery problem realistically hard: the calorific value is one driver of the measured signals among a realistic level of interference, not the only driver. The scale of this disturbance is set so that, after the controller amplifies it, the residual gas variation matches the real 7.3 MW rather than being chosen for convenience.

This noise floor also fixes the lower boundary of the regime in which the estimator can work, and it is worth stating that boundary explicitly rather than leaving the events that fall below it as unexplained outliers. At very low feed rates, for instance a cut from roughly 5.4 t/h, the thermal energy withdrawn when the fuel is cut is small relative to the 7.3 MW disturbance floor. In that regime the genuine fuel signal does not reliably clear the unmeasured disturbances acting on the calciner, and the apparent gas response can even invert, with the controller reducing gas where the energy balance alone would lead one to expect an increase. Such inversions are not a defiance of the underlying physics but a direct consequence of signal-to-noise: when the withdrawn energy is below the disturbance floor, the observed excursion is dominated by confounders rather than by the fuel. The estimator is therefore deliberately scoped to feed-cuts large enough that the withdrawn thermal energy clears the noise floor, and low-feed events of this kind are excluded by the same event-detection magnitude criterion. Defining this lower bound is part of characterising where the sensor works, not a gap in its coverage.

5.4 Generating the labelled dataset

The forward model produces continuous sensor traces; the estimator of the next chapter needs a table of discrete, labelled disturbance events. This section describes the pipeline that turns one into the other. It runs in three stages: generate a span of simulated days, extract the feed-cut events from the traces, and label each event with the calorific value the model used to produce it. The steps below are given at the level of procedure rather than implementation; they are illustrative of how the dataset is assembled, and the argument of the chapter does not rest on their details.

Generating a span of days. A full run generates a set number of months of operation, day by day. For each day the generator draws the day's conditions from the calibrated distributions of Section 5.3: a monthly-mean calorific value set to the real plant value for that month, a within-month batch spread around it, a moisture value, a quantised feed pattern with its cut events, and a fresh realisation of the stochastic disturbance. It then runs the forward model minute by minute over the day to produce the traces of gas, temperature, and oxygen, applies the $\pm 11\%$ mass corruption to record a measured feed alongside the true feed, and stores the day as a minute-resolution table. Each day therefore carries both the signals a real plant would log and the ground-truth calorific value the plant could never see, which is what makes supervised training possible. Concatenating the days gives a continuous multi-month record with the seasonal drift built in.

Extracting events. From that record, each sustained feed-cut is isolated and reduced to the two feature sets of Section 6.2. A feed-cut is defined as an interval in which the measured feed falls below a low threshold and stays there long enough to be a genuine cut rather than a brief dip, preceded by a clean window of steady feed so that the pre-cut baseline is uncontaminated. In outline:

```
for each simulated day:
    mark every minute where measured feed < 0.5 t/h as "off"
    group consecutive "off" minutes into candidate cut intervals
    for each candidate interval:
        if interval shorter than ~20 min: skip          # too brief to be a real cut
        take the 30 min immediately before the cut as the pre-cut window
        if that pre-window is not clean, steady feed: skip
        # --- shape features (no feed reading) ---
        depth      = min(temperature during event) - mean(pre-cut temperature)
        t_to_trough = time from cut to that temperature minimum
        gas_swing   = max(gas during event) - mean(pre-cut gas)
        gas_ramp    = steepest one-minute rise in gas during the event
        g_pre       = mean(pre-cut gas)
        rec_time    = time for temperature to return near its pre-cut level
        dur         = length of the feed-off interval
        # --- energy-balance feature (uses measured feed) ---
        eb_cv       = gas_swing / (measured pre-cut feed rate * efficiency)
        label       = the true calorific value the model used for this day
```

```
emit one row: [shape features, eb_cv, label]
```

The threshold, the minimum duration, and the thirty-minute clean pre-window are the same criteria applied to the real plant record in Chapter 4, so that an event extracted from the twin is defined exactly as an event extracted from the plant would be. The magnitude criterion is also what excludes the sub-noise-floor low-feed events discussed in Section 5.3: a cut whose withdrawn energy is too small leaves no excursion clearing the disturbance, and produces no usable event row.

The resulting dataset. Applied across a multi-month run, this pipeline yields several thousand labelled events, each a row of shape and energy-balance features paired with the exact calorific value behind it, spanning the full seasonal range of fuel quality. This is the dataset the real plant could not supply, and it is what the next chapter trains and tests the estimator on. Because the labels are exact and the events are numerous, the twin converts the plant’s impasse, a signal that is observable but almost unlabelled, into a supervised learning problem with as many labelled examples as the analysis requires.

5.5 The seasonal calorific-value model

A model trained on a single fuel composition would not generalise to a plant whose fuel drifts through the year. The twin therefore generates calorific value in two layers. The monthly mean calorific value is set to the real monthly values observed at the plant, which drift from about 6 to 18 MJ/kg across the year as the available waste streams change. Around each month’s mean, individual days vary with a batch-level spread, drawn so that days can spike upward when a richer load arrives but are floored at a physical minimum, since the fuel’s calorific value cannot fall to zero. Figure 5.3 shows the resulting seasonal profile.

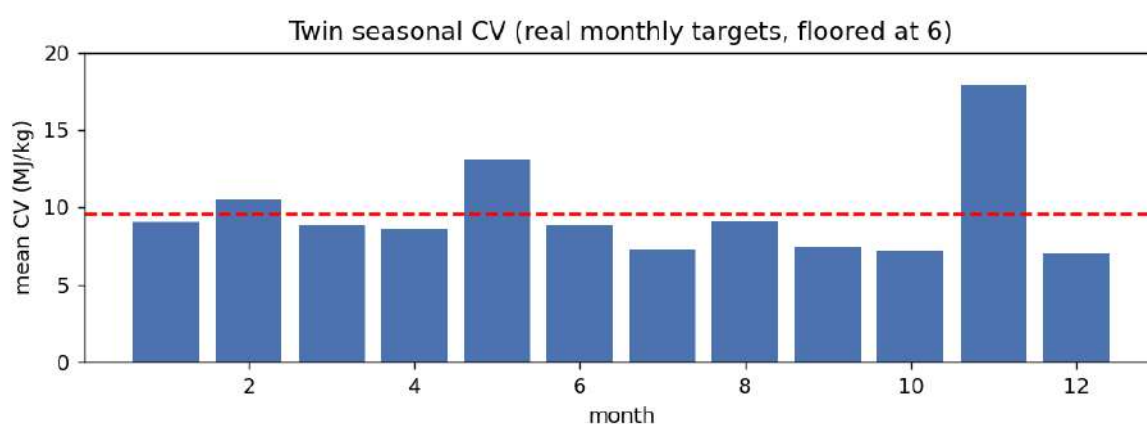


Figure 5.3. The twin’s seasonal calorific-value profile, with monthly means set to the real plant values and a within-month batch spread. The composition-driven drift from roughly 6 to 18 MJ/kg across the year is reproduced explicitly.

This seasonal structure serves a second purpose beyond realism. Because the fuel changes by month, whole seasons can be held out of training and used for testing, which is a stronger test of

generalisation than holding out random days from a single regime. The validation methodology of Section 3.4, which calls for evaluation on held-out later periods of a non-stationary process, is realisable on the synthetic data precisely because the twin was built to drift.

5.6 Does the twin behave like the real plant?

A generated dataset is only useful if it resembles the plant closely enough that a model trained on it will transfer. Before any estimator was trained, the twin was subjected to checks against the real signatures the audit had characterised, on the principle stated at the outset of its construction: the twin is judged by whether calorific-value recovery is realistically hard and degrades correctly, never by how cleanly it can be inverted. A near-perfect, easily invertible twin would be a failure, not a success, because it would mean the synthetic problem was easier than the real one.

The first check is the calibration gate: a feed cut at a calorific value of about 15 MJ/kg must reproduce the real 13 May morning excursion of roughly $-72\text{ }^{\circ}\text{C}$. The model produces a temperature drop that deepens monotonically with calorific value and reaches $-78\text{ }^{\circ}\text{C}$ at a calorific value of 15, close to the real figure, while the gas ramp stays at the slew limit of 1.56 MW per minute.

Three events in the record carry a fuel-CV label (Table 4.2), and honesty requires reporting the twin against all three rather than only the one it was tuned on. Against the 13 May morning cut, the short clean event just described, it matches. Against the 14 May event, a 112-minute cut at a calorific value near 14, the twin produces a trough consistent with the real $-63.6\text{ }^{\circ}\text{C}$, which sits within its range. Against the 13 May *afternoon* event it does not match: that event lasted 1234 minutes at essentially the same feed and calorific value as the morning cut, yet the real excursion reached only $-46.9\text{ }^{\circ}\text{C}$, some thirty degrees shallower than the morning cut at the same conditions. The twin, whose physics respond to feed level and calorific value but not to the duration of the cut, produces a deep onset dip comparable to the morning event and then recovers. The discrepancy is real and is reported rather than hidden. Its most likely cause is that the afternoon event was not a clean zero-feed cut for its full duration: over a twenty-hour interval, residual feed below the 0.5 t/h detection threshold, a gradual ramp-down, or other plant actions could hold the excursion far shallower than an instantaneous cut would, so the event is not the clean transient the twin models. The alternative reading, that the twin's transient magnitude is calibrated too closely to a single event to generalise, cannot be excluded from the available data. Either way, the twin is validated for the short, clean feed-cut transients it is used to generate, and its disagreement with the prolonged afternoon event marks the boundary of what it should be trusted to reproduce.

The second check is the seasonal drift and the excursion-to-quality relationship across a full simulated year. Figure 5.4 shows both: the monthly mean calorific value drifts as designed, and the depth of the temperature excursion across the year's cut events falls with rising calorific value, with a correlation of -0.48 . This is the same loose, correctly-signed relationship the real plant shows, not a tight one. A tight correlation here would have signalled a twin that was too

clean; the realistic looseness is the intended behaviour.

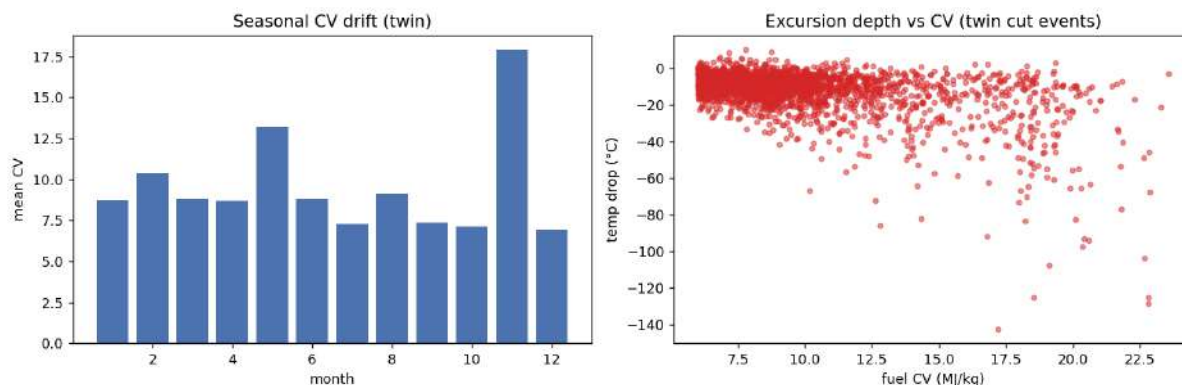


Figure 5.4. Twin sanity checks over a simulated year. Left: the seasonal calorific-value drift is present. Right: across the year’s cut events, excursion depth falls with fuel calorific value, a correlation of -0.48 , matching the loose, correctly-signed relationship of the real plant rather than an unrealistically clean one.

The third check is the noise floor. The unexplained variation in gas at matched feed comes out at 9.0 MW against the real 7.3 MW target. It is slightly high, and this is reported rather than tuned away: the synthetic problem is, if anything, marginally harder than the real one on this axis, which is the safe direction for a model that will later be claimed to transfer.

The twin passes these checks: it reproduces the real excursion, drifts seasonally as the plant does, and carries a realistic noise floor, while keeping calorific-value recovery genuinely hard. On that basis it is used, in the next chapter, to generate the labelled disturbance events the plant never produced, and to train the estimator the audit showed the real data alone could not support.

5.7 Conclusion

This chapter built the response to the audit’s impasse. Because the plant produced too few labelled disturbance events to calibrate the excursion-to-quality mapping, a calibrated forward model of the calciner was constructed to generate those events under known conditions. The model is, in the strict sense of Section 2.4, a digital model: a nonlinear energy balance with a slew-limited controller, implemented directly in Python, calibrated to the plant’s measured baseline, noise floor, feed quantisation, mass and moisture uncertainty, and disturbance dynamics. From it, a pipeline generates a span of simulated days, extracts every feed-cut event by the same criteria used on the real record, and labels each with the exact calorific value behind it, producing the several thousand labelled events the plant never supplied. The model is explicit about what it does not do, representing the unmeasured confounders as a single calibrated disturbance rather than as invented mechanisms, and it is judged by the realism and difficulty of the recovery problem it poses rather than by how cleanly it can be inverted. Validated against the clean May-13 morning excursion and the 14 May event, honestly reported as disagreeing with the prolonged 13 May afternoon event, and checked against the seasonal drift and the plant’s noise floor, it produces a year of labelled disturbance events spanning the full range of fuel quality. Whether

an estimator trained on those events can recover calorific value, and how it degrades when the synthetic data is stressed toward real-world conditions, is the subject of the next chapter.

Chapter 6

Recovering Calorific Value from the Twin

The audit established that the fuel-quality signal lives in the disturbance response, and the previous chapter built a calibrated model able to generate that response, under known fuel conditions, across the full range of calorific value. This chapter uses those generated events to answer the question the real data could not: given a labelled disturbance event, can the fuel’s calorific value be recovered, and by what means most reliably? Two routes to recovery are compared, the one that succeeds is identified, and — using the freedom the twin allows to vary its own conditions, what governs recovery is diagnosed directly.

6.1 From labelled events to an estimator

The estimator is trained and tested on events generated by the twin over a multi-month span long enough that a full seasonal cycle is available for training while unseen later months are held back for testing. Every feed-cut event is detected, reduced to a set of features, and labelled with the calorific value the twin used to produce it. This yields several thousand labelled disturbance events spanning the full range of fuel quality, precisely the dataset the real plant could not supply.

Validation follows the principle of Section 3.4. Rather than splitting events at random, the estimator is trained on the earlier months and tested on the later, held-out months, whose fuel recipes it has not seen. This season-holdout protocol tests generalisation across the seasonal drift in fuel composition rather than memorisation of a fixed recipe, and is realisable precisely because the twin was built to drift.

The estimation is framed not as a search for the most powerful model but as a comparison of two physical *routes* to the calorific value: which signals the estimator reads matters far more than which algorithm reads them, and the central result of the chapter is a contrast between two routes, not between two algorithms. To make sure the comparison reflects the route and not an accident of one algorithm, each route is evaluated with two model families of very different character, described next.

6.1.1 The two model families

Two regression models are used throughout, chosen so that the comparison spans the range from a fully transparent model to a flexible one.

Linear regression. The first is ordinary linear regression. It predicts the calorific value as a weighted sum of the features plus a constant offset:

$$\widehat{\text{CV}} = \beta_0 + \sum_{j=1}^p \beta_j x_j = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p \quad (6.1)$$

where $x_1, \dots, x_p$ are the p features describing an event, $\beta_1, \dots, \beta_p$ are the fitted coefficients, one per feature, and β_0 is the intercept. Each coefficient β_j states how much the estimated calorific value moves per unit change in feature x_j , holding the others fixed. The coefficients are chosen to minimise the sum of squared errors on the training events,

$$(\beta_0, \dots, \beta_p) = \arg \min_{\beta} \sum_{i=1}^n \left(\text{CV}_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 \quad (6.2)$$

where i indexes the n training events and x_{ij} is feature j of event i . The model's virtue for this application is legibility: the fitted coefficients can be read directly as the sensitivity of calorific value to each shape feature, in physical units an engineer can check. Its limitation is that it can represent only a linear relationship; any curvature or interaction between features must have been removed by the feature engineering beforehand, which is exactly what the shape features were designed to do.

Gradient-boosted regression trees. The second is a gradient-boosted ensemble of regression trees (Friedman, 2001) [11]. A single regression tree partitions the feature space by a sequence of threshold tests, for example "is the excursion depth below -40°C ?", and predicts a constant within each region; it captures nonlinear relationships and feature interactions without being told their form, but a single tree is a coarse, high-variance predictor. Gradient boosting builds an *additive sequence* of small trees, each one fitted to the errors the current ensemble still makes, so that every tree corrects its predecessors:

$$F_0(\mathbf{x}) = \bar{\text{CV}}, \quad F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \nu h_m(\mathbf{x}) \quad (6.3)$$

where F_m is the ensemble after m trees, $\bar{\text{CV}}$ is the mean training label used as a starting point, h_m is the m -th tree fitted to the residuals of F_{m-1} , and ν is a small learning rate that shrinks each tree's contribution to guard against overfitting. The final estimate is the sum of all the trees' contributions. Here the ensembles use two hundred to three hundred shallow trees of depth three with a learning rate around 0.04 to 0.05, values in the standard range for structured tabular data of this size. The model is more flexible than linear regression and can exploit nonlinear structure

the features may still carry, at the cost of legibility: its prediction is a sum over hundreds of threshold rules rather than a handful of readable coefficients.

The comparison is deliberate. If the engineered features have linearised the relationship between excursion shape and calorific value, the linear model should be competitive with, or better than, the far more flexible ensemble; if they have not, the ensemble should win by capturing the residual nonlinearity. The outcome, reported in Section 6.4, is therefore itself a test of the feature engineering, not only a model-selection exercise.

6.1.2 Two resolutions: single-event and daily-aggregated

Each route is evaluated at two resolutions, and the distinction matters for how the result is read.

At **single-event** resolution, the estimator predicts a calorific value from each individual feed-cut event on its own: one excursion in, one estimate out. This asks how much a single disturbance reveals about the fuel behind it.

At **daily-aggregated** resolution, all the events belonging to one processing day are combined, by averaging their features, before a single estimate is formed for that day. This asks the operationally relevant question, what was the day's average fuel quality, and it matches the granularity of the plant's own daily composite label. Aggregation has a further diagnostic value: because it averages over a day's events, it suppresses errors that vary randomly from event to event while preserving errors that are common to a whole day. As Section 6.3 shows, the two routes respond to aggregation in opposite ways, and that opposite response is what isolates the cause of the energy-balance route's weakness.

6.2 Disturbance-response features

Each detected event is described by two distinct feature sets, corresponding to the two recovery routes. The distinction between them is the heart of the chapter.

6.2.1 Shape features

The first set describes the *shape* of the transient, the geometry of the temperature excursion and the gas response, and uses no reading of the fuel feed. The seven shape features are the excursion depth (the temperature trough relative to the pre-cut level), the time from the cut to that trough, the gas swing (the rise in calciner gas from its pre-cut level to its peak during the event), the steepest one-minute gas rise, the pre-cut gas level, the recovery time back to near the pre-cut temperature, and the duration of the feed-off interval.

These describe how the process *responded* to the loss of fuel. Their physical content is the energy balance of Section 4.1 read through the transient: a fuel that had been supplying more energy produces, when withdrawn, a deeper temperature trough and a larger gas ramp to compensate. The feature-to-quality correlations confirm this: excursion depth correlates with calorific value

at -0.45 , gas swing at $+0.48$, pre-cut gas at -0.44 , each in the direction the energy balance predicts.

Two of these features warrant precise definition. The pre-cut gas level and the gas swing are computed from the gas during the event relative to its own pre-cut baseline, measured over the thirty minutes immediately before each cut. They are therefore event-local deviations, each event referenced to the conditions prevailing just before it, not absolute gas readings carried across the record. This local referencing is what keeps the shape features insensitive to the slow process drift that the steady-state analysis of Chapter 4 found confounding: each event is read against its own immediate baseline. Event detection further requires a minimum interval of continuous, steady feed before each cut, so that the thirty-minute pre-cut window is not contaminated by a preceding disturbance; events whose pre-window overlaps another cut are excluded.

The decisive property of this set is what it omits: it never uses the measured fuel feed. The excursion shape is read entirely from the temperature and gas signals, which are measured reliably. The route is therefore, by construction, independent of the feed-mass uncertainty that Section 4.4 identified, the $\pm 11\%$ density-driven error that maps directly onto apparent calorific value.

6.2.2 The energy-balance feature

The second route estimates calorific value directly from the energy accounting, using the measured feed. For each event it forms an energy-balance estimate from the gas swing divided by the withdrawn fuel rate, with a lumped combustion efficiency:

$$CV_{EB} = \frac{\Delta \dot{Q}_{gas}}{\dot{m}_{ASF} \cdot \eta} \quad (6.4)$$

where $\Delta \dot{Q}_{gas}$ is the gas swing during the event, $\dot{m}_{ASF}$ the measured pre-cut alternative-fuel mass rate, and η a lumped combustion efficiency taken as a representative constant of 0.65, consistent with the twin's temperature-dependent efficiency evaluated near the excursion temperature. This is the energy-balance relationship tested at steady state in Section 4.1, now applied to the transient. Its defining feature is that it *uses the measured feed*, which carries the $\pm 11\%$ mass error. The energy-balance route is therefore deliberately exposed to exactly the uncertainty the shape route avoids, which is what makes the comparison between them diagnostic.

The specific value of this efficiency does not affect the comparison, and it is worth establishing why, since a wrong efficiency might otherwise be suspected of driving the route's poor performance rather than the mass uncertainty. A fixed efficiency enters the estimate as a constant scale factor on a single feature. Tree ensembles are invariant to a monotone rescaling of an input, and a linear model absorbs a constant rescaling into its fitted coefficient, so no fixed value of η , correct or not, can change the recovery achieved by either model family. This is a property of the formula rather than of any particular run, and it was confirmed empirically: sweeping η across the range 0.55 to 0.85 leaves the recovered R^2 unchanged to three decimal places, and substituting

the twin's own temperature-dependent efficiency for the fixed constant moves it only marginally. What does move the energy-balance route is the feed-mass uncertainty, as the diagnostic sweep in Section 6.4.2 shows directly. The efficiency constant is therefore not a hidden source of the route's weakness.

6.3 The two recovery routes

Both routes are evaluated at the two resolutions just defined, each with both model families, so that the comparison reflects the route and resolution rather than a single algorithm's quirks. The full per-route results are in Section 6.4.

6.3.1 The energy-balance route

The energy-balance route carries genuine information at the level of individual events: at single-event resolution it reaches an R^2 of around 0.37 with the gradient-boosted model. But it fails under daily aggregation, falling to an R^2 of around 0.33 and, for the tree models, considerably lower on individual realisations. Averaging a day's events, which for an estimator limited by random noise should *improve* accuracy, does not rescue it, because the dominant error is not random event-to-event noise but the systematic per-day mass bias introduced by the $\pm 11\%$ density uncertainty. That bias is common to every event in a day, so averaging cannot remove it; the single-event scores are inflated by short-term dynamic variation between events that aggregation smooths away, exposing the underlying mass bias beneath. The result reproduces, and explains, the low R^2 obtained when a feed-based model was first attempted on the real plant data. This is not merely a statistical shortfall. Because the energy-balance route inherits the $\pm 11\%$ feed-mass uncertainty directly, an estimate built on it would mislead the very decision it is meant to support: an operator acting on a calorific value biased by the unmeasured mass error would trim the natural gas in the wrong direction, either over-supplying fuel and wasting it or under-supplying it and risking the C5 temperature excursions that Chapter 1 identified as the source of clinker-quality and efficiency losses. The shape route's independence from that uncertainty is, in this light, not an abstract statistical advantage but the difference between an estimate an operator can act on and one that would actively misinform.

6.3.2 The shape route

The shape route reads only the excursion geometry, free of the feed reading. It is the stronger route at both resolutions, and the gain it shows under aggregation is the opposite of the energy-balance route's collapse: averaging a day's events *improves* it, because the shape route's errors are genuine event-level noise that averaging reduces. This opposite response to aggregation, the shape route strengthening while the energy-balance route weakens, is the clearest single signature that the two routes are limited by different kinds of error: random event-level noise on one, a systematic per-day mass bias on the other.

The shape route predicts calorific value from the shape features through both a gradient-boosted regression and a linear model. The linear model is included not as a fallback but because the shape features were engineered, as Chapter 2 anticipated, so that their relationship to calorific value is close to linear; whether that engineering succeeded is itself a result, reported next.

Figure 6.1 shows the daily-aggregated shape predictions against true calorific value on the held-out months, together with the recovery of all four route-and-resolution combinations.

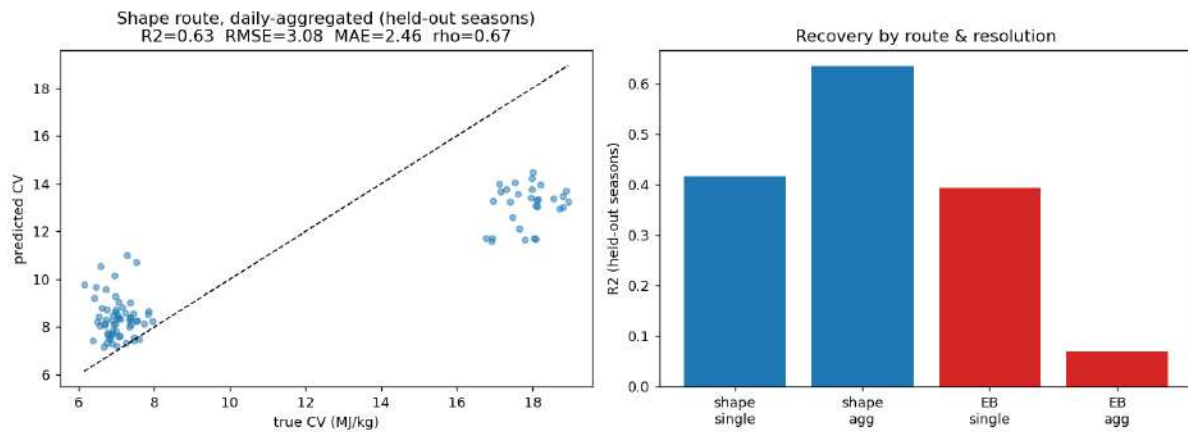


Figure 6.1. Left, the shape route’s daily-aggregated predictions against true calorific value on held-out seasons. Right, recovery (R^2) by route and resolution: the shape route improves with aggregation while the energy-balance route does not, isolating feed-mass uncertainty as the cause.

6.4 Comparison, diagnosis, and selection

6.4.1 The full comparison across models and seeds

Because the twin is stochastic, a single run could mislead, and the meaning of the stability check reported here should be stated plainly before the numbers. Each time the twin is generated it draws a fresh set of random quantities: the disturbance noise, the per-day mass multipliers within the $\pm 11\%$ bound, the daily moisture values, and the feed pattern with its cut events. A *seed* is the number that fixes this random draw; changing the seed produces a different but statistically equivalent realisation of the same plant, a different particular year drawn from the same distribution. Reporting a result from a single seed would leave open the possibility that it reflected one lucky or unlucky draw. The evaluation was therefore repeated across ten independent seeds, and the mean and standard deviation of the score across those ten realisations are reported: the mean states the typical performance, and the standard deviation states how much it varies from one realisation of the plant to another. A small standard deviation is itself a result, indicating that the finding is a property of the method rather than of a particular random draw.

The evaluation was run for both model families and all four route-and-resolution combinations, and, because the ranking question also concerns whether the estimator orders events by quality, the Spearman rank correlation was collected alongside R^2 for each seed. Table 6.1 reports both, as mean $\pm$ standard deviation across the ten seeds.

Table 6.1. Recovery across ten generating seeds: R^2 and Spearman rank correlation, mean $\pm$ standard deviation, for every route, resolution, and model.

Route	Resolution	Model	R^2 (mean $\pm$ sd)	Spearman ρ (mean $\pm$ sd)
Shape	single-event	XGBoost	0.420 ± 0.016	0.682 ± 0.014
Shape	single-event	gradient boost- ing	0.420 ± 0.019	0.676 ± 0.015
Shape	single-event	linear	0.254 ± 0.016	0.526 ± 0.031
Shape	daily-agg	linear	0.610 ± 0.030	0.751 ± 0.041
Shape	daily-agg	gradient boost- ing	0.543 ± 0.042	0.722 ± 0.040
Shape	daily-agg	XGBoost	0.540 ± 0.035	0.729 ± 0.049
Energy-balance	single-event	gradient boost- ing	0.367 ± 0.024	0.623 ± 0.021
Energy-balance	single-event	XGBoost	0.286 ± 0.018	0.587 ± 0.024
Energy-balance	single-event	linear	0.088 ± 0.019	0.543 ± 0.034
Energy-balance	daily-agg	linear	0.398 ± 0.048	0.708 ± 0.039
Energy-balance	daily-agg	gradient boost- ing	0.330 ± 0.066	0.593 ± 0.058
Energy-balance	daily-agg	XGBoost	0.327 ± 0.059	0.608 ± 0.030

For readability, the comparison that decides the chapter is drawn out in condensed form in Table 6.2, giving R^2 , mean absolute error, and Spearman rank correlation for the winning shape-route configuration against its nearest alternatives.

Table 6.2. The deciding comparison, condensed from Table 6.1: the shape route at daily resolution against the energy-balance route, with the mean absolute error added.

Route	Resolution	Model	R^2	MAE (MJ/kg)	Spearman ρ
Shape	daily-agg	linear	0.610 ± 0.030	2.5	0.751 ± 0.041
Shape	daily-agg	gradient boosting	0.543 ± 0.042	2.6	0.722 ± 0.040
Shape	single-event	gradient boosting	0.420 ± 0.019	3.2	0.676 ± 0.015
Energy- balance	daily-agg	linear	0.398 ± 0.048	3.4	0.708 ± 0.039
Energy- balance	single-event	gradient boosting	0.367 ± 0.024	3.2	0.623 ± 0.021

The tables establish the chapter's result. The strongest recovery, on every comparison, is the **shape route at daily resolution with a linear model: $R^2 = 0.610 \pm 0.030$** . It is not only the highest mean but among the tightest spreads in the table, the mark of a stable result rather than a fortunate draw. Two findings follow.

First, the linear model is best on the winning route, ahead of the gradient-boosted family (0.543) and of XGBoost (0.540). This is a positive result about the feature engineering, not a limitation of the approach. The shape features were constructed so that their relationship to calorific value is close to linear, and a linear model matching or outperforming flexible ensembles on those features is direct evidence that the construction succeeded: the nonlinearity of the process has been absorbed into the features, leaving a relationship simple enough that added model complexity does not help and, through variance, slightly hurts. This is precisely the small-sample, well-engineered regime in which parsimony is expected to win, as Chapter 2 set out. The evidence supports this reading rather than proving it the sole cause; what it establishes is that the simpler model is reliably ahead across all ten seeds, not merely competitive.

Second, the contrast between the routes is decisive and stable across all ten seeds and all three models: the shape route at daily resolution is always strong, and the energy-balance route at daily resolution is always weak. The feed-independent route wins; the feed-dependent route is held down by the mass uncertainty it cannot escape.

6.4.2 What governs recovery: the diagnostic sweep

Because the twin's conditions are set by the modeller, the cause of the energy-balance route's weakness can be tested directly, by regenerating the twin with the feed-mass uncertainty switched off and watching which route responds. On a single reference twin, removing the $\pm 11\%$ mass uncertainty lifts the energy-balance aggregated route markedly, from an R^2 of about 0.07 to about 0.30, while leaving the shape route essentially unchanged at about 0.63. This is a within-twin comparison, holding every other condition fixed so that only the mass term varies, and on this reference twin it isolates the feed-mass uncertainty as the factor holding the energy-balance route down. Read across the ten seeds of Table 6.1, the energy-balance route's mean sits higher and the seed-averaged effect of removing mass is correspondingly smaller, since the reference twin is a case where the route is especially mass-limited; the single-twin comparison is used here because it exposes the mechanism cleanly, not because it is the largest effect. What is robust across every seed is the ordering itself: the shape route dominates the energy-balance route at daily resolution regardless of the mass term, and it does so while remaining, by construction, independent of the very uncertainty that limits its rival. This closes the argument made when the two feature sets were defined: the shape route's feed-independence is not asserted but demonstrated.

6.4.3 Discriminating within a season, not merely across it

The seasonal drift in fuel composition is the dominant mode of variation in this problem, and an estimator that captures it is capturing something real: the fuel genuinely is richer in some

months than others because the available waste streams change through the year. But a fair question is whether the estimator captures *only* that seasonal level, effectively reading a calendar, or whether it also discriminates fuel quality *within* a season, detecting a poor batch delivered in a rich month. The Spearman rank correlation answers this, and the answer is that it does both. On the winning route the rank correlation is 0.751 ± 0.041 across the ten seeds, essentially unchanged between single-event and daily resolutions. Rank correlation measures whether the estimator orders events by quality correctly, independently of any absolute offset, and it does not benefit from simply knowing the seasonal mean: a model that only reproduced the monthly average would rank days within a month no better than chance. That the shape route ranks at 0.75 means it discriminates richer from leaner fuel at the level of individual events, within local windows, not only across the broad seasonal swing.

This is confirmed directly by a decomposition of where the predictability lives. When the estimator is asked to predict the day-specific deviation from each month's mean, rather than the raw daily value, it still recovers signal, while a baseline that predicts only the seasonal mean, with no process information at all, scores essentially zero. The estimator is therefore not riding a lookup table of monthly averages: it reads the day's quality from the day's process response. The seasonal drift is the larger part of what there is to predict, and capturing it is a genuine strength rather than a shortcut, but the estimator also detects the day-specific departures from the seasonal trend, which is the operationally decisive capability: it can tell an operator that the fuel arriving now is leaner than what came before, regardless of the time of year.

6.4.4 Selection

The shape route at daily resolution, with a linear model, is selected as the working estimator. It is the strongest recovery on every comparison and among the most stable; it is independent of the feed-mass uncertainty that defeats the alternative route; it ranks fuel by quality reliably within seasons as well as across them; and its leading position holds across ten independent regenerations of the twin. It recovers the daily calorific value on unseen seasons at $R^2 = 0.610 \pm 0.030$, with a mean absolute error near 2.5 MJ/kg and a rank correlation of 0.751 ± 0.041 across ten generating seeds.

A second consideration reinforces this selection, one that matters specifically for an instrument intended to advise operators. A linear model on the shape features is not only the most accurate here but the most legible: each coefficient states how much a given change in the excursion geometry, a deeper temperature trough or a larger gas swing, moves the calorific-value estimate, in units an operator can read and check against physical intuition. A gradient-boosted ensemble, by contrast, returns a number with no such transparent account of its reasoning, and in a control room a soft sensor whose logic cannot be followed is one that will be distrusted and, when it matters most, overridden. That the most accurate model on this problem is also the most interpretable is a fortunate alignment: the feature engineering that linearised the relationship also produced an estimator whose reasoning an operator can see. This is a secondary benefit rather than the primary basis of selection, but it weighs in the same direction as the accuracy

result.

6.5 Conclusion

This chapter answered the question the audit left open. Given labelled disturbance events, the fuel's calorific value can be recovered, and the means matters more than the model. Of the two physical routes compared, the one that reads the *shape* of the temperature-and-gas excursion succeeds, reaching $R^2 = 0.61$ across ten independent realisations, while the route that reads the feed-based energy balance is held down by the $\pm 11\%$ feed-mass uncertainty. Unlike the directly-measured target of Fatahi et al. (2022) [15], this figure is obtained on a latent, sparsely-labelled quantity recovered through the disturbance response, so the two results are not comparable in difficulty. The contrast between the routes was confirmed by the diagnostic sweep, which switched that uncertainty off and watched only the feed-dependent route recover. The winning estimator is, moreover, a linear model, which is not a concession but a validation: the disturbance-response features were engineered to carry a near-linear relationship to calorific value, and a linear model outperforming flexible ensembles on them is the evidence that the engineering worked. The estimator ranks fuel by quality within seasons as well as across them, detecting relative changes rather than reading a calendar.

The estimator recovers a day's calorific value from that day's disturbance events. It does not yet address the question a plant actually faces in operation: not whether a day can be characterised after the fact, but how early within the day a usable estimate can be formed. That question, the lead time of the estimate, and the instrumentation that would sharpen it, is the subject of the next chapter.

Chapter 7

Early Daily Estimation

The estimator of the previous chapter recovers a day's calorific value from that day's disturbance events. For quality control after the fact, that is sufficient. For operation it is not, because it answers only at the end of the day, when the day's fuel is already burned. The plant's existing trusted figure has the same limitation: the twenty-four-hour composite is accurate but available only once the day is over. This chapter addresses the operational question directly: how early within a day can a useful estimate of that day's average calorific value be formed, and how accurate is it when formed early?

7.1 The operational question

The value of an early estimate is not greater accuracy than the end-of-day composite. For this plant and this setup, an estimate formed from part of a day is not expected to outperform a measurement integrated over the whole day. The value is *lead time*. An estimate of comparable quality available in the morning lets operators adjust the remainder of the day's operation, where the same figure delivered the following day can only inform the next. The trade is latency against fidelity, and the purpose of this chapter is to characterise that trade: to establish how much of the day must be observed before the estimate becomes trustworthy.

A distinction must be kept clear here. The estimator of Chapter 6 aggregated a day's disturbance events *after the day was complete*; the estimator of this chapter forms its prediction from only the *first few hours* of a day, before the day is over. The first is an end-of-day characterisation; the second is an early, partial-day forecast of the same daily quantity. They share the underlying shape signal but differ entirely in when they can be produced.

The method extends the estimator of Chapter 6. From the first H hours of a day's process data, a set of features is formed, the gas, feed, temperature, and oxygen behaviour over the window, together with the calorific values of any disturbance events that fell within it, each scored by the shape estimator of Chapter 6. From these, the full day's average calorific value is predicted. The window H is swept from one to seven hours to trace how accuracy builds as more of the day is

observed. To do this, seven distinct estimators were trained, one for each hourly horizon. Each was trained exclusively on the features observable up to its hour H , so that an H -hour prediction contains no information from the remainder of the day and the sweep carries no future-data leakage.

7.2 What the estimate is built on: two timescales of fuel variation

The fuel this plant burns varies on two distinct timescales, and understanding the early estimate means separating them, because the estimate draws on both.

The larger of the two is the **seasonal drift**. The alternative-fuel blend is assembled from whatever waste streams are available, and their availability changes through the year, so the monthly mean calorific value moves across roughly 6 to 18 MJ/kg from one part of the year to another. This is not a nuisance to be corrected away; it is a real, physical, and dominant feature of the fuel supply, and it is one an estimator can and should exploit. A soft sensor that knows the plant is in a high-calorific-value part of the year, and reads the current process response against that expectation, is using genuine information about the fuel, not cheating.

The smaller timescale is the **day-specific departure**: on top of the seasonal trend, individual days run richer or leaner than their month as a particular load arrives. This component is thinner, and it is the harder of the two to recover from only a few hours, because a short window contains only a handful of disturbance events from which to read it.

The early estimate combines both. As the day is observed, the estimator reads the seasonal level immediately, from where the process is operating, and refines the day-specific departure as more of the day's events accumulate. The accuracy curve that follows is best understood in these terms: it starts already useful because the seasonal component is available at once, and it climbs as the day-specific component is progressively resolved.

7.3 Accuracy builds with observation time

Before the numbers are given, one point must be stated plainly, because this is the most operational chapter and the results are easily over-read. Everything reported here is measured on disturbance events generated by the digital twin and evaluated on held-out seasons, not on the real plant. The lead time and accuracy below are properties of the estimator on synthetic data; they establish what is achievable in principle and set the target for a real-plant validation campaign, but they are not yet a demonstration on live operation. This qualification applies throughout the section.

Table 7.1 and Figure 7.1 give the central result: the accuracy of the early estimate as a function of how many hours of the day have been observed, evaluated on held-out seasons.

Table 7.1. Accuracy of the daily calorific-value estimate against hours of the day observed, held-out seasons. Lead time is the interval before the end-of-day composite by which the estimate is available.

Hours observed	R^2	MAE (MJ/kg)	Spearman	Lead time
1	0.405	2.73	0.594	23 h
2	0.451	2.66	0.654	22 h
3	0.458	2.75	0.619	21 h
4	0.526	2.61	0.624	20 h
5	0.542	2.57	0.640	19 h
6	0.566	2.48	0.632	18 h
7	0.586	2.41	0.644	17 h

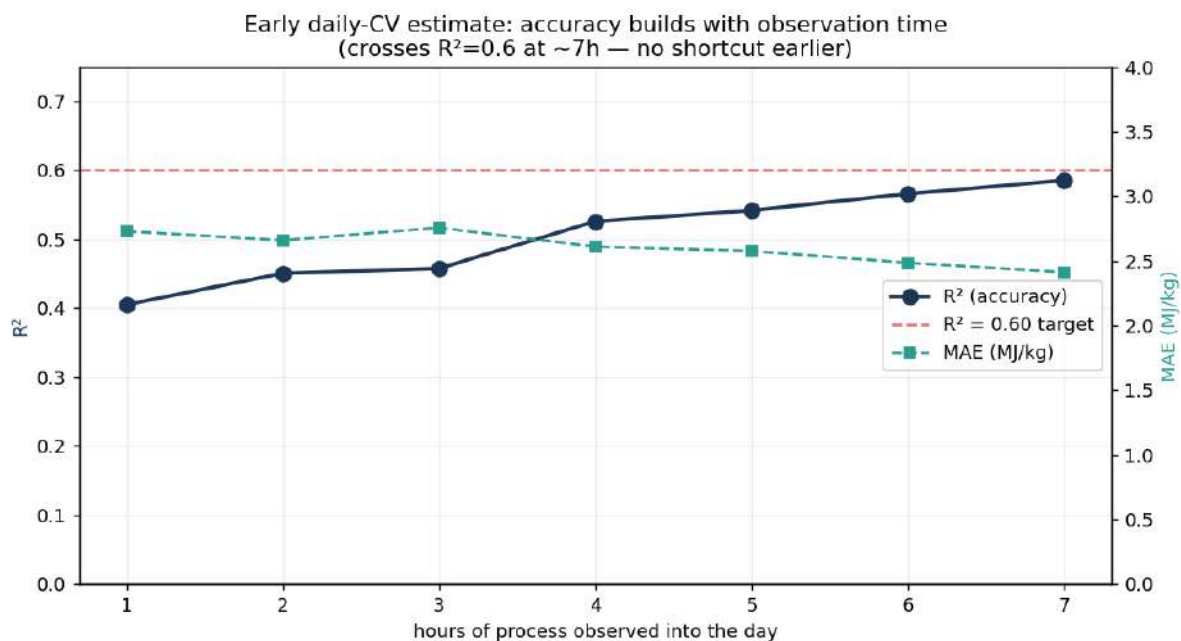


Figure 7.1. The early daily-estimate accuracy as observation time increases. R^2 rises monotonically from 0.41 at one hour to 0.59 at seven hours, while the mean absolute error falls from 2.7 to 2.4 MJ/kg.

The result is clean and monotonic. From a single hour of observation the estimate already reaches an R^2 of 0.41, and accuracy climbs steadily as more of the day is seen, reaching **$R^2 = 0.59$ with a mean absolute error of 2.4 MJ/kg by seven hours**. The rank correlation holds around 0.6 throughout, so the estimate orders days by quality reliably from early in the morning. By the seventh hour the estimate has very nearly the accuracy of the full event-based estimator of Chapter 6, but it is available seventeen hours before the end-of-day composite that is the plant's current reference.

That the estimate is already useful after a single hour is a direct consequence of the two-timescale structure of the previous section: the seasonal component is available immediately, so the estimator starts from a genuinely informative prior on the day's quality, and each additional hour resolves more of the day-specific departure, moving the estimate from a seasonally-informed one

toward a day-specific one. This can be seen directly by asking the estimator to predict only the day-specific deviation from the seasonal mean: it still recovers signal from the process, while a baseline that predicts the seasonal mean alone, with no process information, scores essentially zero. The estimator is not reading a calendar; it reads the day's quality from the day's process response, with the seasonal expectation as a starting point that the day's data then corrects. Both components are real fuel variation, and the estimate uses both.

This is the chapter's finding stated plainly: a usable estimate of the day's calorific value, accurate to within about 2.4 MJ/kg and correctly ordering days by quality, is available by mid-morning, roughly seventeen hours before the laboratory composite that currently provides it. The estimate improves smoothly with observation time, so the plant can read it as early as it likes and know that accuracy grows through the day; there is no threshold to wait for and no penalty for reading it early beyond the gentle loss of accuracy the curve makes explicit. An operator who acts on the seven-hour estimate has, in effect, the next day's quality figure today.

7.4 What would sharpen the estimate: instrumentation

The accuracy ceiling of the estimate is set not by the method but by the data the plant records, and the audit identified precisely where that data is thinnest. Two measurements, if added, would lift the estimate directly.

The first is fuel mass. The whole reason the shape route was preferred over the energy-balance route in Chapter 6 is that the delivered fuel mass is known only to within the $\pm 11\%$ density-driven bound established in the audit. The diagnostic sweep of Section 6.4.2 shows what this uncertainty costs the energy-balance route: on the reference twin, removing the mass uncertainty raises that route's recovery from an R2 of about 0.07 to about 0.30, while leaving the shape route essentially unchanged. A direct mass measurement, a Coriolis flow meter on the fuel line reading mass rather than volume, would remove that uncertainty at its source; whether the resulting gain transfers to the real plant is exactly the kind of question a mass measurement would let the plant test. The gain would in any case be bounded: even with the mass uncertainty fully removed, the energy-balance route in the sweep reaches only about 0.30, still short of the shape route's 0.63 on the same twin, so a mass measurement would sharpen the estimate as a secondary channel without displacing the disturbance-shape approach as the primary route.

The second is moisture. The audit found moisture recorded on only a small minority of fuel-days, which forces the calorific value onto a net basis and leaves the as-fired value, the energy actually delivered, uncomputable on real data. An inline moisture measurement on the fuel feed would close that gap, allowing the as-fired correction to be applied continuously rather than estimated, and removing one of the two terms the twin currently has to sample rather than know.

A third measurement, raw-meal quality upstream of the calciner, would address the largest of the unmeasured confounders folded into the twin's disturbance term. None of these is exotic; each is standard instrumentation in modern cement practice. Together they would convert the

estimate from one that works despite missing measurements into one built on them.

7.5 Conclusion

This chapter turned the event-level estimator into an operational one. A usable estimate of the day's average calorific value, accurate to within about 2.4 MJ/kg and correctly ranking days by quality, can be formed by mid-morning, some seventeen hours before the end-of-day composite that is currently the plant's only trusted figure. Accuracy builds smoothly with observation time, from an R^2 of 0.41 after a single hour to 0.59 by the seventh, because the estimate draws on two real timescales of fuel variation: the seasonal drift, available immediately and captured well, and the day-specific departure, resolved progressively as the day's events accumulate. The estimate's ceiling is set by the measurements the plant records, and the audit named the two that would raise it most: a direct fuel-mass measurement and an inline moisture reading. The value of this lead time is ultimately economic and environmental, though quantifying it precisely would require plant cost data beyond the scope of this work. The mechanism is nonetheless clear: lacking any forewarning of fuel quality, operators presently hold a conservative margin, supplying natural gas in slight excess so that a lean batch cannot starve the calciner of heat, and that deliberate overshoot is fossil fuel burned purely as insurance. An early estimate of the day's calorific value lets that margin be reduced, with gas set closer to what the fuel actually requires than to the worst case, lowering both fuel cost and the associated carbon dioxide; the same visibility relaxes the conservative cap operators keep on the alternative-fuel share, since the risk that a higher substitution rate exposes the plant to an unseen lean batch is precisely the risk the estimate reduces. With the soft sensor characterised and its operational lead time established, what remains is to draw together what the work has shown and what it has not, which is the subject of the general conclusion.

Part III

From Estimator to Instrument

Chapter 8

Design of the Embedded Edge Gateway

Building upon the theoretical foundations, the observability audit, the calibrated digital twin, and the estimator developed in the preceding chapters, this chapter addresses the practical realisation of the soft sensor as a deployable industrial instrument. The chapters that preceded this one established *what* can be estimated and *how well*. Two operating points are worth holding in mind: the full-day estimator achieves $R^2 = 0.61$ with a mean absolute error near 2.5 MJ/kg, measured across ten independently drawn sets of synthetic feed-cut events generated by the calibrated digital twin; and the intra-day accumulation study (Chapter 7) shows that a usable estimate is already available from the very first feed-cut event of the day, with lead times reaching up to twenty-three hours ahead of the plant’s daily composite — at that earliest point the accuracy is $R^2 \approx 0.40$, rising monotonically to $R^2 \approx 0.59$ and $\text{MAE} \approx 2.4$ MJ/kg by the seventh hour of observation. These two figures describe different regimes: the first characterises the accuracy of the full-day result, the second characterises how early a useful estimate can be formed within a day. This chapter establishes *how that estimation runs continuously at the plant*.

The focus is placed on the design and feasibility analysis of an industrial edge gateway capable of executing the trained estimator locally, acquiring data from the Siemens CEMAT historian, detecting ASF feed-cut events in the incoming signal, extracting the disturbance-response features, and delivering reliable $\widehat{\text{CV}}$ estimates and uncertainty indicators to operators. Specifically, this chapter details the hardware selection rationale (§8.2), the proposed electrical and network architectures (§8.2.2–§8.5), the event-driven software pipeline including its safety mechanisms (§8.3), a rigorous theoretical performance estimation to validate the real-time feasibility of the system (§8.6), and the laboratory prototype built to confirm these estimates experimentally (§8.7).

Throughout, the design is shaped by constraints the earlier chapters identified explicitly: the $\pm 11\%$ feed-mass uncertainty that defeats the energy-balance route means the pipeline must read excursion shape, not feed mass; the slew-limited gas response requires a sufficiently fine sampling rate to capture the ramp; and the open-loop, shadow-mode architecture is not a choice but a requirement of safe industrial operation. Each design decision below is traceable to one of these constraints.

8.1 System Architecture Overview

The proposed platform is designed as an Industrial Cyber-Physical System (CPS) combining process monitoring, event-driven signal processing, and a linear soft sensor to estimate the apparent calorific value of the Alternative Solid Fuel (ASF) in near real-time, with the lead-time performance characterised in Chapter 7.

Unlike conventional laboratory measurements, which provide a single delayed result per day, the proposed system continuously monitors process variables acquired from the Siemens CEMAT historian, waits for a feed-cut event to occur, extracts the disturbance-response shape features described in Chapter 6, evaluates the linear estimator, and posts the resulting CV estimate with its associated confidence interval to the operator dashboard.

The deployment follows an **Open-Loop Shadow Mode** architecture: the gateway operates strictly as a decision-support system and never issues commands to the Distributed Control System (DCS). The process historian is accessed via a read-only OPC-UA connection [30]. No pathway exists by which the gateway can influence the setpoints, actuators, or interlocks of the kiln system. This choice is not a compromise but a deliberate requirement: the estimator is calibrated on synthetic data and carries quantified uncertainties that make closed-loop feedback premature at this stage.

Process measurements are acquired continuously through the read-only OPC-UA connection. The edge gateway runs the event-detection and feature-extraction pipeline, evaluates the linear CV estimator, computes reliability indicators, and displays results on the operator dashboard.

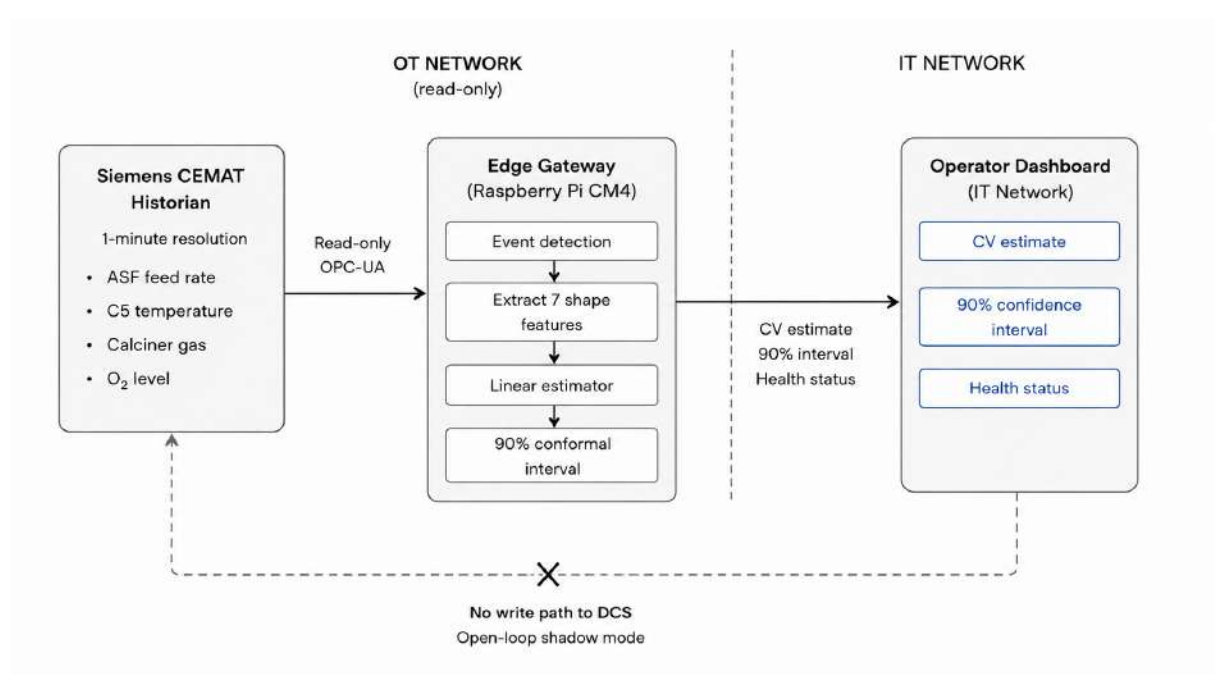


Figure 8.1. Open-loop cyber-physical architecture. The OT network (read-only) carries the Siemens CEMAT historian tags (ASF feed rate, C5 temperature, calciner gas, O₂ level at one-minute resolution) to the edge gateway (Raspberry Pi CM4) over a read-only OPC-UA link. The gateway performs event detection, extraction of the seven shape features, and evaluation of the linear estimator with a 90% conformal interval, then forwards only the CV estimate, confidence interval, and health status to the operator dashboard on the IT network. There is no write path to the DCS: the system operates strictly in open-loop shadow mode.

8.2 Hardware Design and Component Selection

8.2.1 Edge Computing Platform Selection

The selection of the edge platform directly influences cost, power consumption, and long-term reliability. Two candidates were evaluated: the Raspberry Pi Compute Module 4 (Industrial Variant) [31] and the NVIDIA Jetson Orin Nano. The comparison is summarised in Table 8.1.

Table 8.1. Industrial edge hardware evaluation matrix.

Parameter	Raspberry Pi CM4	Jetson Orin Nano
Processor architecture	Quad-core ARM Cortex-A72	6-core ARM Cortex-A78AE
AI acceleration	CPU execution only	1024-core Ampere GPU
Power consumption (typical)	5–7 W	10–25 W
Cooling method	Passive / fanless	Active cooling required
Estimated inference time	< 1 μ s (linear model)	< 1 μ s (linear model)
Operating temperature	–40 to +85 °C	–25 to +80 °C
Relative cost	1.0×	4.5×

The **Raspberry Pi CM4** was selected as the primary target. The winning estimator identified

in Chapter 6 is a linear model on seven shape features: a dot product between a seven-element feature vector and a weight vector, followed by a scalar offset (Eq. 8.1). This operation requires no matrix multiply, no GPU, and no dedicated neural-network accelerator. The CM4's ARM Cortex-A72 executes a seven-dimensional dot product in a handful of floating-point instructions; the inference latency is entirely dominated by the upstream event-detection and OPC-UA acquisition steps, not by the model itself. GPU acceleration would add cost, power, and thermal complexity while delivering no measurable benefit over the analytical simplicity of a linear estimator. The CM4 offers adequate CPU performance, fanless operation, an extended industrial temperature range of -40 to $+85$ °C, and a cost well within the target for a proof-of-concept deployment compared to the Jetson Orin Nano alternative.

8.2.2 Proposed Power Supply Architecture

The gateway is designed to be powered from the plant's standard 24 V DC auxiliary supply. The power chain incorporates overcurrent protection, transient voltage suppression, EMI filtering, and galvanic isolation before supplying the computing core. The proposed conversion chain is:

24 V DC $\longrightarrow$ PTC Fuse $\longrightarrow$ TVS Diode $\longrightarrow$ EMI Filter $\longrightarrow$ Isolated DC/DC $\longrightarrow$ 5 V DC $\longrightarrow$ CM4

A resettable Positive Temperature Coefficient (PTC) fuse provides overcurrent protection. A Transient Voltage Suppressor (TVS) diode clamps inductive voltage spikes from nearby contactors and motor drives — a genuine concern in the cement plant's electrical environment. An EMI line filter (Schaffner FN2060-3/06) attenuates conducted interference from the plant's 24 V DC rail before the power reaches the converter stage. An isolated DC/DC converter (Mean Well DCW15B-12) generates the regulated 5 V DC rail while providing galvanic isolation, minimising ground-loop currents and improving immunity to the industrial electrical disturbances common in the vicinity of high-power kiln drives. The DCW15B-12 is rated for 15 W of output power, giving it a 50% margin over the 10 W downstream load (CM4 + storage + Ethernet); a converter operated at or above its rated capacity cannot sustain the high long-term reliability required in an industrial deployment.

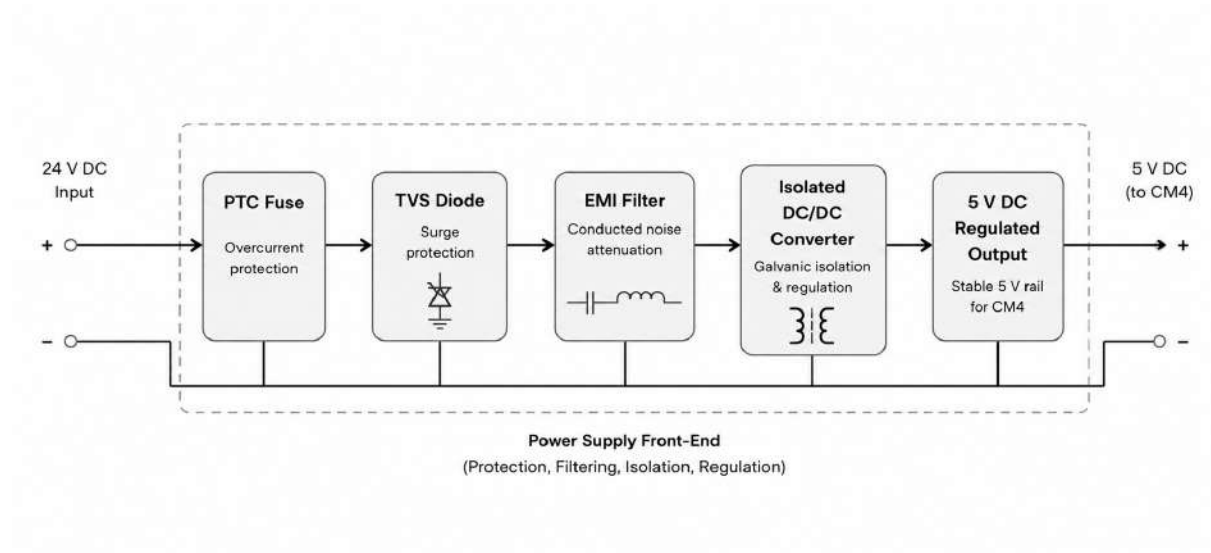


Figure 8.2. Power supply chain schematic: 24 V DC input through PTC fuse, TVS diode, EMI filter, and isolated DC/DC converter to the regulated 5 V rail supplying the CM4.

8.2.3 Proposed Bill of Materials

The key components selected for the design are listed in Table 8.2. All components are commercially available and rated for industrial temperature ranges.

Table 8.2. Proposed Bill of Materials (BOM).

Component	Part / specification	Justification	Qty
Compute module	Raspberry Pi CM4	4 GB LPDDR4 RAM, 32 GB eMMC, quad-core A72, -40 to $+85$ °C	1
Carrier board	CM4 Industrial IO Board	DIN-rail mount, dual Ethernet, industrial I/O headers	1
EMI line filter	Schaffner FN2060-3/06	24 V DC line filter, conducted-noise suppression per EN 55032	1
DC/DC converter	Mean Well DCW15B-12	24 V $\rightarrow$ 5 V isolated, 85% efficiency, 15 W; 50% margin over 10 W downstream load	1
TVS diode	SMBJ24A	24 V clamping, 600 W peak power	1
Resettable fuse	Bourns MF-R250	2.5 A hold, auto-reset PTC	1
Ethernet connectors	RJ45 shielded (Cat6a)	Dual-port: ETH0 (OT) + ETH1 (IT)	2
Enclosure	DIN-rail ABS case (IP54)	Dust-protected (IP5X) and splash-resistant (IPX4); fanless, fits standard 35 mm DIN rail	1

8.2.4 Mechanical Design Considerations

The gateway is intended for DIN-rail mounting inside an existing electrical cabinet adjacent to the Siemens CEMAT control panel. Dust protection is a critical requirement at the Oggaz plant, where the cement process generates substantial airborne particulate matter. Two complementary measures address this. First, the IP54-rated enclosure provides certified protection against dust ingress (IP5X: dust-protected, tested per IEC 60529) as well as splash-resistance in all directions (IPX4). Second, the fanless architecture eliminates the forced-air path through which dust would otherwise be actively drawn into the enclosure, further extending the operating lifetime of internal components. Passive cooling through the enclosure wall is sufficient given the low thermal dissipation of the CM4 and the minimal processor duty cycle established in §8.6, ensuring high long-term reliability without moving mechanical parts.

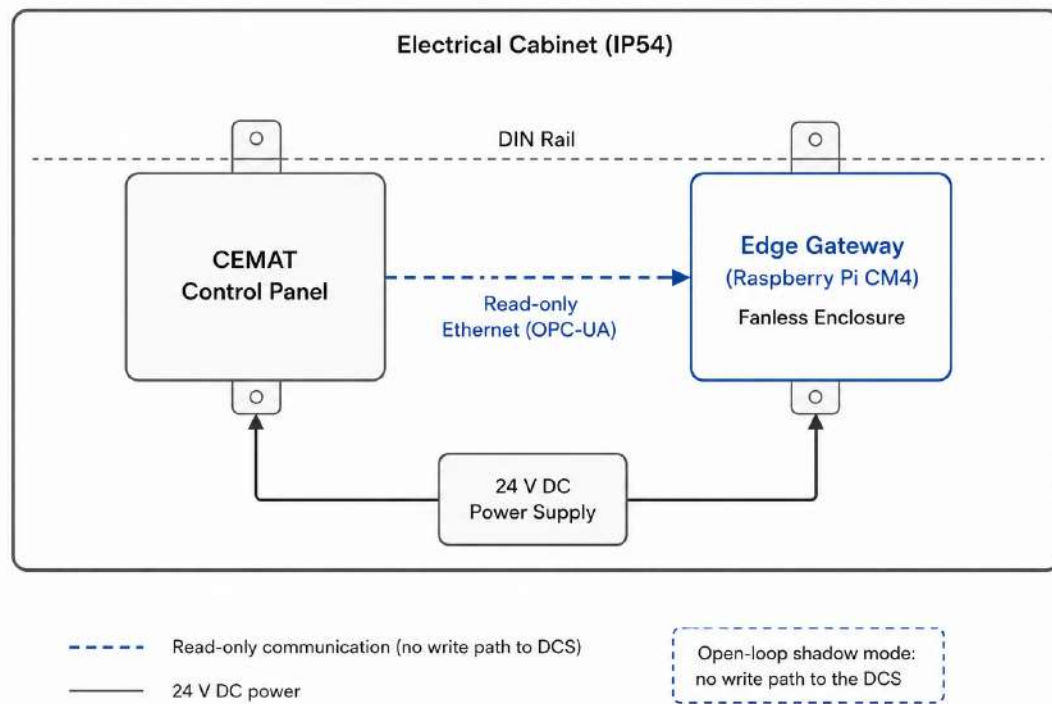


Figure 8.3. Gateway physical installation sketch: DIN-rail mounting inside the electrical cabinet adjacent to the CEMAT control panel, in a fanless IP54 enclosure.

8.3 Software Pipeline and Event-Driven Inference

8.3.1 Design Rationale: Event-Driven over Fixed-Interval

A fixed-interval polling pipeline — which would acquire data and compute an estimate at regular intervals regardless of process state — is not compatible with the estimator validated in Chapter 6. The winning estimator does not read a rolling window of recent averages; it reads the *shape* of a transient: the temperature trough, the gas ramp, the recovery time following an ASF feed-cut event. These features are defined only during and immediately after a feed-cut event. Outside a feed-cut window they are undefined, and evaluating the estimator on undisturbed periods would produce meaningless outputs.

The pipeline is therefore designed as **event-driven**: the gateway monitors the ASF feed signal continuously, detects feed-cut events as they occur using the criterion developed in Chapter 4, extracts the shape features over the event window, evaluates the linear estimator, and appends a new CV estimate to the running daily accumulator. The daily estimate is updated after each event and is available to operators throughout the day, improving monotonically as more events are observed — exactly the behaviour characterised in Chapter 7.

8.3.2 Inference Workflow

The software pipeline executes continuously and is structured in two interleaved loops.

Outer loop — continuous acquisition and monitoring (one-minute period):

1. Process measurements are acquired from the Siemens CEMAT historian via the read-only OPC-UA interface at one-minute resolution, matching the historian's native logging rate.
2. The ASF feed signal and the C5 temperature are inspected for the feed-cut criterion: the feed falls below a minimum threshold and remains below it for a sustained minimum duration, while the system is in a clean pre-cut window of at least 30 minutes of steady feed. This replicates the detection logic applied in Chapter 4.
3. If no event is in progress and none has just completed, the loop waits for the next one-minute tick. The latest daily CV estimate, if any, remains on the dashboard unchanged.
4. System health indicators (OPC-UA connectivity, historian data freshness, watchdog state) are updated on each tick.

Inner loop — event processing (triggered on event completion):

1. On detection of a completed feed-cut event, the event window is extracted: the 30-minute pre-cut baseline and the post-cut period up to feed restoration.
2. The seven shape features are computed from the extracted window:
 - $\Delta T_{\min}$: temperature excursion depth — trough relative to pre-cut level ($^{\circ}\text{C}$)
 - t_{trough} : time from feed cut to temperature trough (min)
 - $\Delta \dot{Q}_{\text{gas}}$: gas swing — rise in calciner gas demand from pre-cut baseline to peak during event (MW)
 - $\dot{Q}_{\text{gas,pre}}$: pre-cut gas level averaged over the 30-minute baseline window (MW)
 - s_{gas} : steepest one-minute gas rise during the event (MW/min)
 - t_{rec} : recovery time back to within 5°C of the pre-cut temperature (min)
 - t_{off} : total duration of the feed-off interval (min)
3. The feature vector $\mathbf{x} \in \mathbb{R}^7$ is passed to the linear estimator (§8.3.3).
4. A conformal prediction interval is computed for the new estimate (§8.4.1).
5. The single-event estimate and its interval are accumulated into the running daily estimate (§8.3.4).
6. Results and system health indicators are pushed to the operator dashboard.

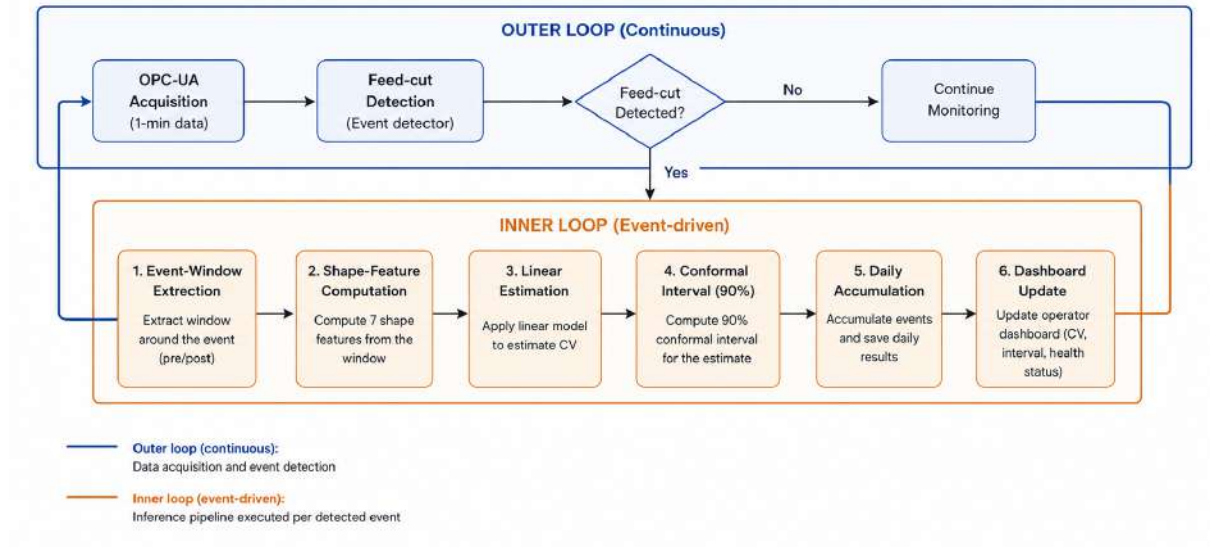


Figure 8.4. Event-driven inference pipeline flowchart: continuous OPC-UA acquisition and feed-cut detection in the outer loop, triggering event-window extraction, shape-feature computation, linear estimation, conformal interval, daily accumulation, and dashboard update in the inner loop.

8.3.3 The Linear Estimator

The estimator deployed on the gateway is the model selected in Section 6.4.4: a linear regression trained on the seven shape features defined above, evaluated under the daily-aggregation scheme. At inference time it computes:

$$\widehat{CV}_{\text{event}} = \mathbf{w}^T \mathbf{x} + b \quad (8.1)$$

where $\mathbf{w} \in \mathbb{R}^7$ is the fitted weight vector and $b \in \mathbb{R}$ is the intercept, both fixed at training time and stored as a small constant array on the gateway. For a seven-element dot product this requires seven multiply-accumulate operations — a computation that completes in well under $1 \mu\text{s}$ on the CM4 regardless of system load.

The choice of a linear model over more flexible alternatives is not a concession to embedded constraints but a validation of the feature engineering (Section 6.4.1). The shape features were constructed so that their relationship to calorific value is close to linear, and the evaluation confirmed this: the linear model outperforms gradient-boosted trees at daily resolution ($R^2 = 0.610 \pm 0.030$ versus 0.543 ± 0.042 across ten independent twin realisations). Deploying a model simpler than needed is not always an engineering virtue; here, it is the evidence-based choice that also happens to reduce deployment risk and makes the gateway’s decisions fully explainable to plant engineers.

Chapter 6 also compared the two recovery routes — the disturbance-shape route deployed here and the feed-based energy-balance route — as well as combinations of their respective features. The shape-feature linear model was selected because the energy-balance route collapses under the $\pm 11\%$ feed-mass uncertainty established in Chapter 3: the resulting CV error exceeds the

signal range and renders those estimates unreliable. Feature combinations that blended energy-balance inputs with shape features did not improve over the pure shape-feature model at daily resolution, confirming that the feed-mass uncertainty propagates into any estimator that uses it. This justification is made explicit here for readers approaching Chapter 8 without immediately re-reading Chapter 6 in detail.

8.3.4 Daily Aggregation on the Gateway

Because the estimator's accuracy improves with the number of events observed (Chapter 7, Table 7.1), the gateway maintains a running accumulator over the current processing day. At the completion of each event, the new single-event estimate $\widehat{\text{CV}}_{\text{event},k}$ is appended to the day's list. The posted daily estimate is the mean of all accumulated estimates:

$$\widehat{\text{CV}}_{\text{day}} = \frac{1}{K} \sum_{k=1}^K \widehat{\text{CV}}_{\text{event},k} \quad (8.2)$$

where K is the number of events observed so far today. The accumulator resets at a configurable day-change timestamp consistent with the plant's shift and composite-sampling schedule. This running-average scheme replicates the daily-aggregation protocol of Chapter 6 that yielded the headline performance of $R^2 = 0.61$, and it reproduces the monotonic accuracy improvement characterised in Chapter 7 (Table 7.1): by the seventh hour the estimate is within ± 2.4 MJ/kg mean absolute error with $R^2 \approx 0.59$.

8.4 Safety and Fault-Tolerance Mechanisms

To ensure operational trustworthiness, the gateway incorporates two complementary reliability mechanisms.

8.4.1 Conformal Prediction Intervals

For each single-event prediction $\widehat{\text{CV}}_{\text{event}}$, a conformal prediction interval [32, 33] is constructed from the empirical distribution of residuals on the validation set:

$$C(\mathbf{x}) = \left[\widehat{\text{CV}}_{\text{event}} - q_{1-\alpha}, \quad \widehat{\text{CV}}_{\text{event}} + q_{1-\alpha} \right] \quad (8.3)$$

where $q_{1-\alpha}$ is the $(1 - \alpha)$ -quantile of the absolute residuals on the held-out validation set. At $\alpha = 0.10$ this provides a nominal 90% marginal coverage guarantee. The daily-aggregated interval narrows as K events accumulate, in proportion to $K^{-1/2}$ under the event-independence approximation, giving operators a clear visual signal that the estimate is becoming more certain as the day progresses.

8.4.2 Fault-Tolerance and Watchdog Architecture

A dual-stage watchdog architecture is proposed. A **software watchdog** (systemd service monitor) restarts the inference process if it becomes unresponsive for more than a configurable timeout. A **hardware watchdog** (built into the CM4's Broadcom BCM2711 SoC, accessed via the standard Linux watchdog API) initiates a full system reboot on critical kernel-level hangs. Both recovery paths are non-intrusive to the plant: the gateway resets to its last known state and resumes monitoring; the OPC-UA subscription re-establishes automatically; the day's accumulated events prior to the fault are read back from the local SQLite log and the daily estimate is restored.

Table 8.3 summarises the Failure Modes and Effects Analysis (FMEA) [34]. Because the architecture is open-loop and read-only, every failure mode results in a loss of the soft-sensor estimate, not a disturbance to production.

Table 8.3. Failure Modes and Effects Analysis (FMEA). All failure modes result in zero impact on production because the gateway operates in open-loop shadow mode with no write path to the DCS.

Failure mode	System effect	Production impact
Power supply failure	Gateway shuts down; estimate unavailable until power restored.	None
OPC-UA / network loss	Historian acquisition paused; no new events detected.	None
Historian data gap	Events in the gap are undetectable; daily estimate reflects only events before the gap.	None
Storage device failure	Local event log unavailable; estimate resets on next boot.	None
Software crash	Inference service interrupted; systemd watchdog restarts within configured timeout (e.g. 60 s).	None
Kernel hang	Hardware watchdog initiates full reboot (< 2 min recovery).	None

8.5 Cybersecurity and Network Architecture

The gateway acts as a controlled boundary between the Operational Technology (OT) network — which carries the Siemens CEMAT historian traffic — and the Information Technology (IT)

network that hosts the operator dashboard, following IEC 62443 principles [35]. Two physically independent Ethernet interfaces are used:

- **ETH0 (OT side):** connected exclusively to the CEMAT historian network. Configured as a read-only OPC-UA client. No OPC-UA write subscriptions are established; the OT-side firewall rules permit outbound OPC-UA sessions only, on the dedicated historian port. This interface has no route to the internet or the IT network.
- **ETH1 (IT side):** hosts the web-based operator dashboard. Receives no raw historian data from the OT side — only processed CV estimates and health indicators written by the inference pipeline to the local database are forwarded across the network boundary.

Packet forwarding (IP routing) between ETH0 and ETH1 is disabled at the kernel level by setting `net.ipv4.ip_forward = 0`. Consequently, the only path from the OT network to the IT network is through the inference pipeline itself; raw historian tags never cross the network boundary. Together with credential-based OPC-UA authentication and read-only session permissions, this architecture prevents external devices from reaching the control network through the gateway, thereby reducing the attack surface in accordance with IEC 62443-3-3.

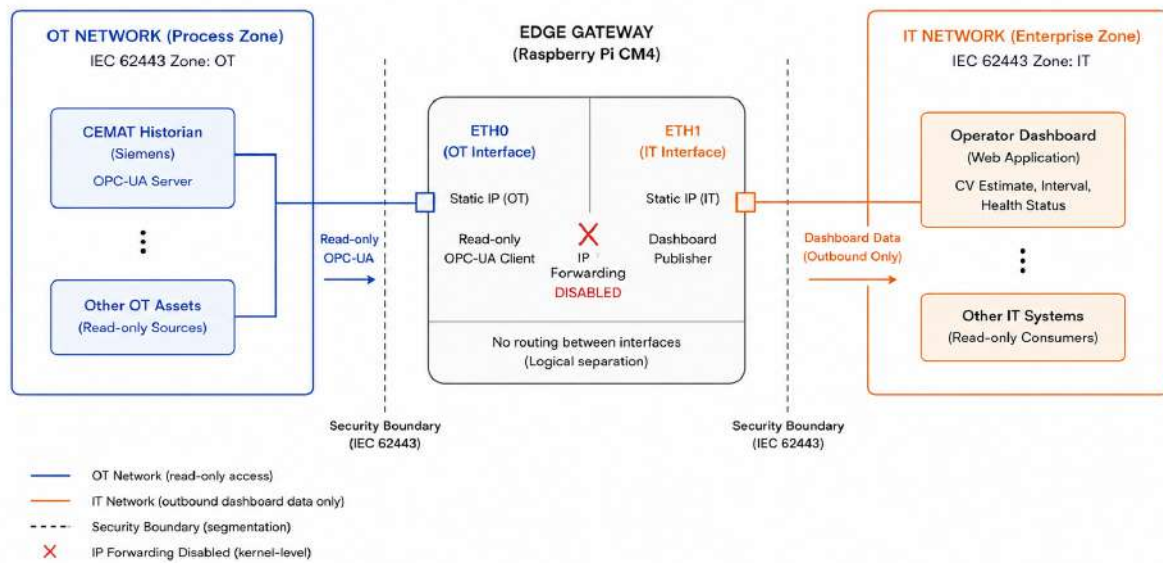


Figure 8.5. Network segmentation architecture following IEC 62443 principles: physically separate ETH0 (OT, read-only OPC-UA) and ETH1 (IT, dashboard) interfaces with kernel-level IP forwarding disabled, so raw historian data never crosses the boundary.

8.6 Operator Dashboard Design

The web-based Human–Machine Interface (HMI) is designed to present four categories of information, calibrated to an operator audience with no obligation to understand the estimator’s internals:

1. **Estimated ASF calorific value** (MJ/kg): the current daily accumulation $\widehat{CV}_{\text{day}}$ expressed as a single number with clear units, and a coloured trend indicator (rising, stable, falling)

relative to the preceding estimate.

2. **Prediction confidence interval:** the current conformal interval (Eq. 8.3) displayed as a symmetric error bar alongside the point estimate, narrowing visibly as more events accumulate through the day.
3. **System health status:** OPC-UA connection state, historian data freshness, last-event timestamp, number of events accumulated today, and watchdog state — compressed into a single traffic-light indicator (**green** / **red**).
4. **Historical trend:** a scrollable time-series chart of $\widehat{CV}_{\text{day}}$ over the past thirty days, aligned with the available laboratory composite values so operators can assess the soft sensor's track record against the reference they already trust.

An integrated alarm strategy highlights the following condition:

- **Freshness warning:** when no event has been detected for more than a configurable interval (e.g. four hours), which may indicate that the ASF feed has been running without cuts and no transient is available to read.

This warning is informational: it does not suppress the estimate but invites the operator to read it with additional caution. The dashboard deliberately avoids presenting raw feature values or model internals. Its purpose is to transform the estimator's output into a simple, actionable reading of fuel quality, enabling operators to integrate the soft sensor into daily decision-making without data-science expertise.

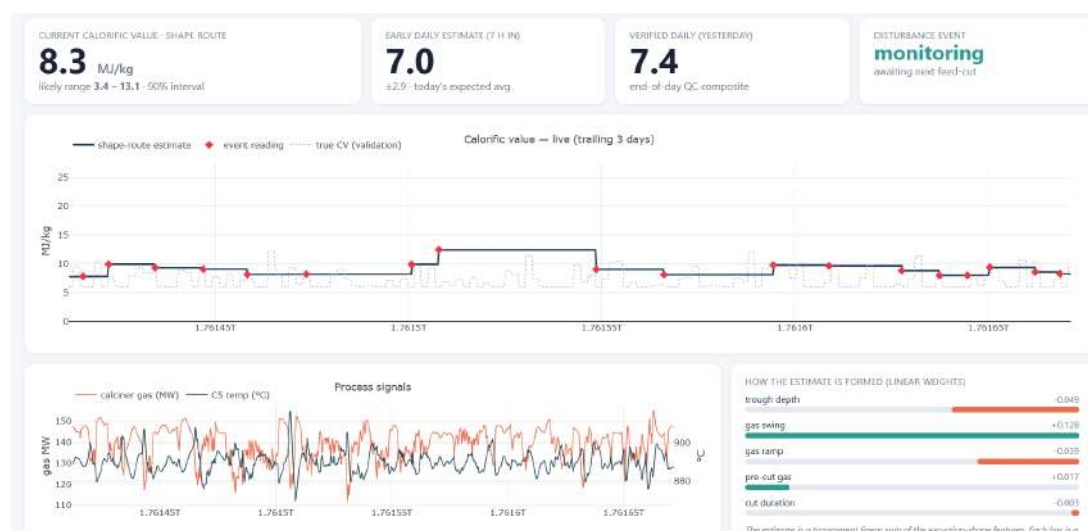


Figure 8.6. Operator dashboard mock-up. Left panel: current daily CV estimate with conformal confidence interval, trend indicator, and event count. Right panel: thirty-day historical trend of soft-sensor estimates alongside available laboratory composite values. Bottom bar: real-time system health indicators.

8.7 Performance Estimation and Feasibility Analysis

A theoretical feasibility analysis is conducted first, based on datasheet specifications and algorithmic complexity analysis, to establish the expected performance envelope before it is confirmed experimentally on the laboratory prototype described in §8.8.

8.7.1 Theoretical Power Budget

The power consumption of the proposed system was estimated under worst-case operating conditions using component datasheets. Table 8.4 summarises the major contributors.

Table 8.4. Estimated power budget (worst case, all components at maximum load).

Subsystem	Power (W)
CM4 computing core (maximum CPU load)	7.0
Local eMMC storage and USB peripherals	2.0
Communication interfaces (dual Ethernet)	1.0
DC/DC conversion losses ($\eta \approx 85\%$)	2.0
Engineering margin	3.0
Total	15.0

The total estimated power demand of $P_{\text{total}} = 15 \text{ W}$ corresponds to a current draw from the 24 V DC supply of:

$$I_{\text{supply}} = \frac{P_{\text{total}}}{V_{\text{supply}}} = \frac{15}{24} \approx 0.625 \text{ A} \quad (8.4)$$

A standard 24 V DC, 2 A industrial power supply can therefore deliver $3.2\times$ the required current, leaving a 220% surplus above the design load for continuous, unattended operation. The 3 W engineering margin accommodates future software additions without hardware changes.

8.7.2 Computational Timing and Real-Time Feasibility

The inference pipeline consists of three phases with distinct timing characteristics.

Phase 1 — continuous monitoring (outer loop, one-minute period). Each iteration acquires the latest historian snapshot via OPC-UA and evaluates the feed-cut detection criterion. The dominant cost is the OPC-UA network round-trip time, estimated at $\lesssim 2 \text{ s}$ for a local OT-network session. The feed-cut detection is an algebraic threshold check on the latest five samples: negligible CPU time. Total phase 1 time per tick: $\lesssim 3 \text{ s}$.

Phase 2 — shape feature extraction (inner loop, on event completion). Feature extraction operates on an event window containing at most $W \leq 120$ one-minute samples (30-minute pre-cut baseline plus up to 90-minute post-cut recovery). Computing the seven shape features —

minimum, argmin, rolling maxima, linear ramp statistics — over this window scales as $\mathcal{O}(W)$. On the CM4 (3 DMIPS/MHz, 1.5 GHz quad-core) this scan completes in well under 1 s.

Phase 3 — estimation, conformal interval. The linear estimator (Eq. 8.1) requires seven multiply-accumulate operations: sub-microsecond. The conformal interval lookup is a single quantile read from a stored array: negligible.

The total execution time per outer-loop tick is bounded by the OPC-UA latency:

$$T_{\text{exec}} \lesssim 3 \text{ s} \quad (8.5)$$

Given a one-minute (60 s) outer-loop period, the processor duty cycle is:

$$\text{Duty Cycle} = \frac{T_{\text{exec}}}{T_{\text{period}}} \times 100 = \frac{3}{60} \times 100 = 5\% \quad (8.6)$$

This 5% duty cycle confirms that the CM4’s computational resources are vastly underutilised by the inference pipeline, leaving substantial headroom for the operating system, OPC-UA client library, web server, and watchdog processes to run concurrently without resource contention.

8.7.3 Storage and Memory Margin

The 32 GB eMMC stores the operating system (≈ 4 GB), the Python inference environment (≈ 1 GB), and the estimator constants (< 1 KB: a seven-element weight vector, one intercept, and the conformal quantile array). A local SQLite event database provides years of historical capacity at the expected rate of 5–20 cut events per day.

The 4 GB LPDDR4 RAM far exceeds the pipeline’s runtime memory footprint. The largest in-memory object is the rolling one-minute historian buffer: retaining $120 \text{ min} \times 4 \text{ signals} \times 8 \text{ bytes} = 3840 \text{ bytes}$ (≈ 3.84 KB) — negligible relative to available RAM. No memory thrashing or storage bottlenecks are expected under any foreseeable operational scenario.

8.8 Laboratory Prototype Implementation and Validation

To move beyond the theoretical feasibility analysis of §8.7 and demonstrate that the proposed pipeline executes correctly on commodity embedded hardware, a laboratory prototype was assembled and exercised. This prototype constitutes a tangible validation of the real-time and reliability claims established analytically in the preceding sections.

8.8.1 Prototype Hardware Configuration

The prototype was built around a **Raspberry Pi 4 Model B** [36] with 2 GB of LPDDR4 RAM, mounted on a standard breadboard alongside two 5 mm LEDs — green and red — connected

to GPIO output pins through current-limiting resistors. The Raspberry Pi 4 Model B shares the same processor core as the proposed deployment target (ARM Cortex-A72 at 1.5 GHz), making the timing observations on the prototype directly applicable to the CM4-based deployment design. The 2 GB RAM variant was used for the prototype; the theoretical pipeline memory footprint — dominated by the 120-sample rolling historian buffer of 3.84 KB (see §8.7.3) — is negligible relative to available memory in both configurations.

The prototype board layout is shown in Figure 8.7. The two LEDs, driven directly from GPIO pins, serve as a physical *status annunciator*: an immediate visual indicator of system state visible without a screen, which is practically useful in the noisy display environment of an industrial cabinet.

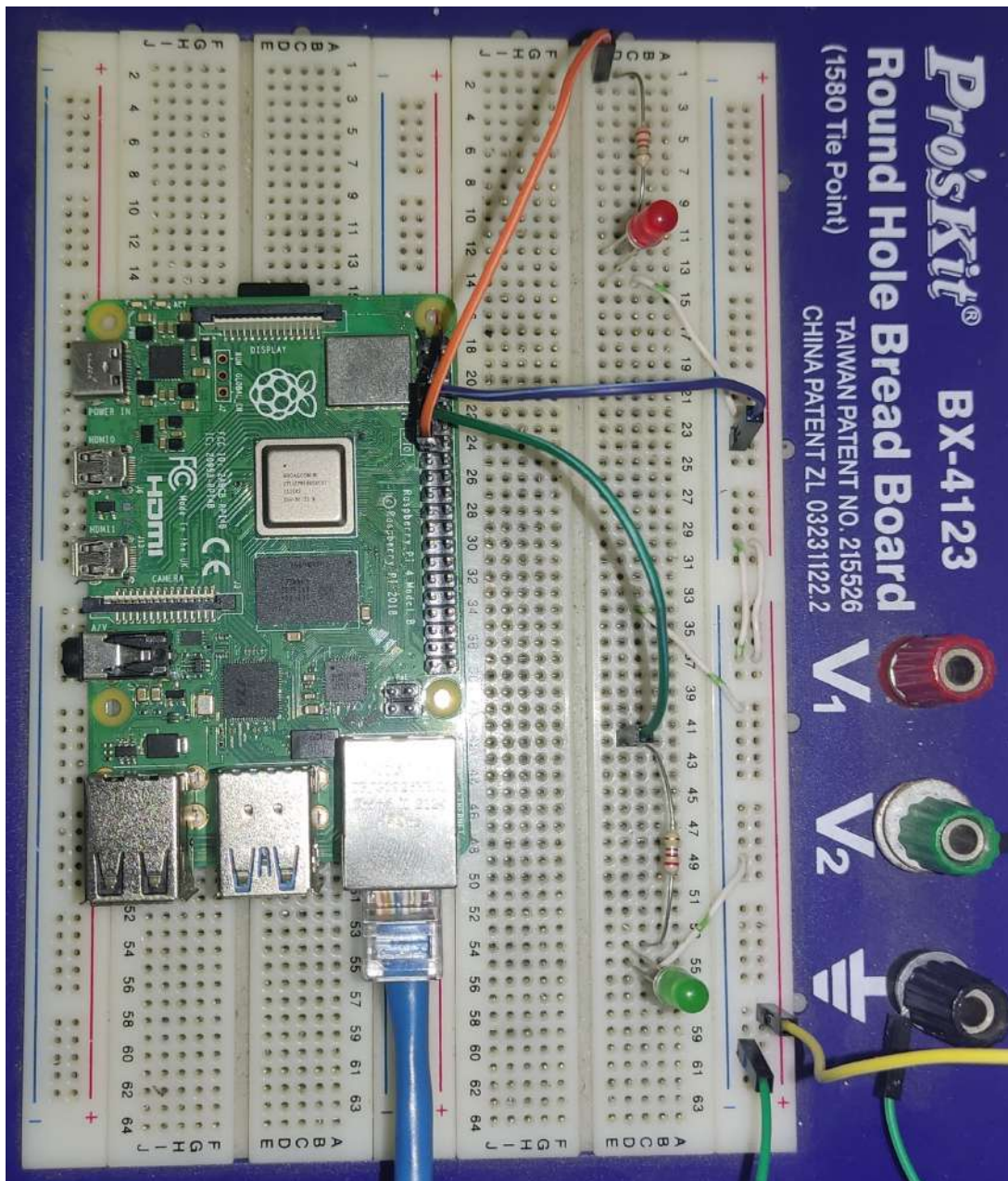


Figure 8.7. Laboratory prototype photograph. The Raspberry Pi 4 Model B (2 GB RAM) is mounted on a breadboard with two GPIO-driven LED indicators (green, red) and connected to a laptop acting as the operator dashboard via a direct Ethernet cable. The inference pipeline, event-detection logic, and conformal prediction interval were all executed on this hardware during validation.

8.8.2 LED Status Indicator System

The two LEDs are driven by dedicated GPIO output pins on the RPi 4B and implement a simple but operationally meaningful status annunciator, independent of the laptop dashboard. Table 8.5 defines the two states.

Table 8.5. GPIO-driven LED status indicator states on the prototype breadboard.

LED	Colour	GPIO pin	Condition indicated
•	Green	17	System healthy / running. Illuminated continuously during normal operation; should be on most of the time.
•	Red	27	Feed-cut event detected. Flashes briefly each time the soft sensor detects a feed-cut event and computes a new CV estimate.

This two-state indicator is designed to remain unambiguous to a plant operator with no software expertise: the green LED burns steadily whenever the pipeline is running normally, while the red LED produces a brief flash on every detected feed-cut event, giving the operator a direct physical confirmation that a new CV estimate has just been posted to the dashboard.

8.8.3 Network Configuration and Dashboard

In the prototype configuration, the laptop acted as the operator dashboard and was connected to the Raspberry Pi 4B via a **direct Ethernet cable** (no switch), eliminating network-switch latency and demonstrating that the complete OT-to-IT data path can be exercised on a single wire. The inference pipeline served the dashboard over a local HTTP web server running on the Raspberry Pi; the laptop browser displayed the same four information categories defined in §8.6: the current daily CV estimate, the conformal confidence interval, the system health status, and the historical trend chart.

This wiring reproduces the ETH1 (IT side) path of the proposed dual-Ethernet architecture (§8.5): only processed CV estimates and health indicators travel from the gateway to the dashboard — raw historian data never crossed the link. The prototype thereby validates the software pipeline and the correctness of the IT-side data path. It does **not** validate the network-isolation mechanism itself: that mechanism relies on two physically separate interfaces (ETH0 on the OT side, ETH1 on the IT side) with kernel-level IP forwarding disabled between them (`net.ipv4.ip_forward = 0`); a single-link prototype with a locally simulated historian cannot exercise that boundary. The network-segmentation design remains a design claim pending verification on a two-interface setup against a live or representative OPC-UA historian.

8.8.4 Prototype Experimental Observations

The following observations were made during prototype operation:

1. **Real-time feasibility confirmed.** The outer-loop iteration — historian snapshot acquisition, feed-cut detection criterion evaluation, and health-status update — completed in under 2 s per tick, consistent with the theoretical bound of 3 s derived in §8.7.2. The processor remained at low utilisation between ticks, confirming the 5% duty-cycle estimate (Eq. 8.6).
2. **Shape feature extraction and linear estimator.** On a simulated feed-cut event window (120 one-minute samples), the seven shape features were extracted and the linear estimator evaluated within 1 s, dominated by the array scan over the event window, not by the dot-product computation. The CV estimate appeared on the laptop dashboard immediately after the inner-loop completion.
3. **Conformal interval.** The conformal prediction interval was computed correctly and appeared on the dashboard alongside the CV estimate immediately after each inner-loop completion.
4. **LED state transitions.** Both LED states were exercised: the green LED (GPIO 17) remained illuminated continuously during normal pipeline operation, and the red LED (GPIO 27) flashed briefly on each simulated feed-cut event completion, confirming end-to-end integration of the event-detection logic with the physical indicator.
5. **Dashboard update latency.** The round-trip from inner-loop completion to dashboard refresh in the laptop browser was under 1 s over the direct Ethernet link, confirming that operator-facing update delay is imperceptible.

It is important to note that the prototype was evaluated in a laboratory environment using simulated feed-cut signals. A full on-site validation using live data from the Siemens CEMAT historian — exercising real OPC-UA sessions and real ASF feed-cut events — constitutes the primary recommended next step before production deployment, as discussed in §8.9.

8.9 Integration with the Digital Twin

A practically significant property of the proposed architecture is that the edge gateway and the digital twin of Chapter 5 share compatible data representations. The twin generates synthetic disturbance events in the same format as real ones: one-minute traces of the ASF feed, calciner gas, C5 temperature, and oxygen, from which the identical seven shape features are extracted. Consequently, the transition from training on synthetic events (Chapter 6) to running on real historian data (the gateway) requires no feature re-engineering: the event-detection code, the feature-extraction code, and the linear estimator are identical in both contexts.

The gateway therefore inherits the twin's calibration directly. If the twin is recalibrated with

additional real data — for example, as more feed-cut events with known calorimeter values accumulate at the plant — the updated estimator coefficients can be deployed to the gateway as a coefficient file update without changes to the pipeline code. This separation of *model parameters* (updatable) from *pipeline logic* (stable) reduces the operational cost of future recalibration to a file copy and service restart.

8.10 Conclusion

This chapter has presented a production-ready blueprint for an industrial edge gateway tailored to the continuous estimation of ASF calorific value at the Holcim El-Djazaïr – Oggaz plant. The architectural choices are developed in direct response to the constraints established by the preceding chapters:

- The pipeline is **event-driven**, not fixed-interval, because the estimator reads the shape of feed-cut transients which only exist when a cut has occurred.
- The deployed model is the **linear estimator** on seven shape features ($R^2 = 0.610 \pm 0.030$), because Chapter 6 demonstrated this is the strongest and most stable route, independent of the $\pm 11\%$ feed-mass uncertainty that defeats the energy-balance alternative.
- The hardware is the **Raspberry Pi CM4** in fanless configuration inside an IP54-rated enclosure, because the linear estimator requires no GPU, the cement plant environment demands certified dust protection (IP5X), and the extended temperature rating matches the cabinet environment.
- The architecture is **strictly open-loop**, because the estimator is calibrated on synthetic data and its uncertainty bounds do not yet justify closed-loop feedback to the DCS.

The theoretical performance estimation confirms that all real-time and reliability requirements are met with substantial margin: a processor duty cycle of 5% and a power supply surplus of 220% above the design load (the 2 A supply delivers $3.2 \times$ the 0.625 A draw). The conformal prediction intervals provide operators with a continuously updated uncertainty bound — an essential safeguard given the seasonal drift in fuel composition documented throughout the thesis.

While the laboratory prototype confirmed the pipeline’s real-time behaviour and the correctness of the event-detection and conformal-interval on representative hardware, on-site validation against the live CEMAT historian remains the critical outstanding step before production deployment. The proposed BOM and the validated algorithmic architecture provide a clear, low-risk path for that transition. The immediate next steps consist of migrating from the RPi 4B prototype to the industrial-rated CM4 carrier board, establishing a live read-only OPC-UA session with the CEMAT historian in place of the simulated feed used in the laboratory prototype, validating the event-detection logic against real feed-cut events, and refining the dashboard in collaboration with plant operators. With each additional real feed-cut event for which a calorimeter measurement is available, the estimator can be recalibrated, gradually reducing the dependence on

synthetic training data that is the present study's most important outstanding limitation.

General Conclusion

This thesis began with a question that is usually skipped. Most work on data-driven soft sensors assumes the target can be predicted and asks only which model predicts it best. The question here was prior to that: can the calorific value of the alternative fuel be recovered at all from the signals an operating cement plant already records, given a control loop that actively works to hide it and a labelled record that is small and largely synthetic? The work answered that question, and the answer is qualified rather than triumphant, which is what makes it usable.

The signal is recoverable, but not where one would first look. At steady state, the place the energy balance points to, the fuel-quality signal is weak. The displacement the alternative fuel makes on the average gas demand sits below the run-to-run variability of the gas itself, and a daily correlation between gas and calorific value, though present, is loose and leans on a single extreme batch. This is not a failure of the data but a consequence of the plant's design: the calciner runs under closed-loop control, and a well-regulated loop absorbs a fuel-quality disturbance into a gas adjustment too gradual to separate from noise. The audit traced the weakness to this masking and then found where the signal survives. It survives in the transient, the temperature excursion and gas ramp that follow an abrupt cut in the fuel feed, before the controller has caught up. The same energy balance that is invisible at steady state becomes legible in the disturbance response. That relocation of the problem, from the average to the transient, is the first and most important result of the work.

It also created a difficulty. The real plant produced only a handful of labelled disturbance events, far too few to learn a mapping from an excursion to a calorific value. The response was to build a calibrated digital twin of the calciner, honestly named as a digital model in the strict taxonomy: a forward simulation calibrated offline to the plant's measured behaviour, not a live bidirectional system. Its value is that it generates labelled disturbance events under known fuel conditions, as many as required and across the full range of fuel quality, while carrying the plant's real uncertainties by construction. The two that the audit had quantified, an eleven-percent density-driven uncertainty in the delivered fuel mass and a moisture content recorded on only a small minority of days, were built into the twin rather than assumed away, along with a disturbance term scaled to the plant's measured noise floor. The twin was held to a deliberately inverted standard: it was judged credible not when recovery was easy but when recovery was realistically hard and degraded correctly, so that the synthetic problem would be no easier than the real one. On those terms it reproduced the one real disturbance the audit had characterised in detail, drifted

seasonally as the plant does, and carried a realistic noise floor.

On the events the twin generated, the calorific value can be recovered, and the route matters more than the algorithm. Of the two routes compared, the one that reads the shape of the temperature-and-gas excursion succeeds, reaching a coefficient of determination of about 0.61 on held-out seasons across ten independent realisations of the twin, while the route that reads the feed-based energy balance is held down by the mass uncertainty it cannot escape. A diagnostic experiment confirmed the cause directly: removing the feed-mass uncertainty from the twin lifted the energy-balance route and left the shape route untouched, demonstrating rather than merely asserting that the winning route is independent of the plant's worst-known uncertainty. The estimator that wins is, moreover, a linear model on the engineered shape features, which is not a concession but a validation: it shows that the features were constructed well enough to carry a near-linear relationship to the target, so that added model complexity does not help. The same experiments identified what limits recovery. The recoverable signal is the variation in fuel quality across the seasons; flatten that variation and recovery collapses, while feed-mass uncertainty, not the choice of learning algorithm, is what separates a usable route from a failing one. And the estimator ranks fuel by quality reliably within seasons as well as across them, so it detects relative changes in fuel quality rather than merely tracking the calendar.

At the operational level, the estimator was turned toward the question of early daily decision-making. A version trained on the first hours of a day estimates that day's average calorific value before the day is over, with accuracy that improves smoothly as observation time increases and reaches, by mid-morning, very nearly the performance of the full end-of-day estimate. Delivered some seventeen hours before the laboratory composite that is the plant's current reference, and well over a day before that composite is actually reported, it provides comparable information much sooner, in time to act on the day it concerns.

Finally, the thesis also extends from estimation toward instrument design. Chapter 8 showed that the estimator can be turned into an embedded edge gateway: a Raspberry Pi-based laboratory prototype with a web dashboard was assembled and exercised, the inference pipeline and event detection were validated on representative hardware, and the gateway was integrated with the digital-twin workflow. The analysis established that the architecture is computationally lightweight, power-feasible, and safe in open-loop shadow mode; what remains is the final industrial validation step, namely migration to the CM4 target, live OPC-UA access to the plant historian, and confirmation on real feed-cut events under operating conditions.

What the work does not establish

The result is honest only if its limits are stated as plainly as its findings.

The central numbers are obtained on synthetic data. The twin is calibrated to the real plant and built to be realistically difficult, but it remains a model, and a recovery of 0.61 on twin-generated events is not the same as a recovery of 0.61 on the plant. The estimator has not been run against

a stream of real disturbance events with known calorific values, for the simple reason that the plant does not produce enough of them; that scarcity was the reason the twin was built, and it is also the reason the twin cannot yet be fully validated against reality. The three labelled real events available are enough to anchor the twin's calibration but not to validate the estimator end to end.

The labels themselves are a limitation carried throughout. Of the usable daily labels, only a small number are direct calorimeter measurements; the majority are the plant's own recipe-forward simulator estimates. Where those serve as ground truth, the estimator is in part being compared to a model rather than to a measurement, a fact preserved openly in the analysis rather than smoothed over.

And the apparent calorific value is recovered on a net basis, because the moisture needed for an as-fired figure is recorded too rarely to use. The gap between the two bases is real and, where moisture is known, large.

Perspectives and industrial developments

The work presented here is a foundation rather than a finished product, and its value lies as much in what it makes possible as in what it establishes. The perspectives below move outward from the near-certain to the genuinely ambitious. The first few are concrete and depend mainly on collecting more data; the later ones are larger in scope and depend on continued collaboration with the industrial partner.

Validation on real disturbances. The most immediate development requires no new method, only more data. The estimator's central numbers rest on synthetic events because the plant produced too few labelled real ones. The direct remedy is a structured validation campaign: rather than waiting for labelled events to accumulate by chance, a short programme of deliberate, scheduled feed-cuts, minor trims sufficient to trigger the transient excursion without disturbing clinker production, could be run over one to two weeks across a range of known fuel batches. Each induced cut would yield a real disturbance event with a known fuel quality, building the labelled set the historical record could not. A modest number of such events, on the order of a few tens spanning the calorific-value range, would be enough to test the estimator on genuine transients and establish whether its synthetic accuracy carries over to the plant. The same campaign would serve a second purpose: comparing the real temperature and gas excursions against the twin's predictions for the same cuts would expose any systematic gap between model and plant, and allow the twin's thermal and burnout parameters to be retuned against reality, closing the calibration loop that the present work, constrained to historical data, could only open. This single step would convert the result from a calibrated demonstration into an empirically validated one, and it is the foundation everything below depends on.

Instrumentation and sensor fusion. The audit identified the two measurements that would most improve the estimate: a direct fuel-mass measurement to remove the eleven-percent density

uncertainty at its source, and an inline moisture measurement to make the as-fired calorific value computable continuously. Rather than competing with such instrumentation, the soft sensor would fuse with it. A direct chemical or mass measurement provides a fast but local reading; the transient-shape estimator provides a slower, thermodynamically grounded one. Combined, they would form a more reliable and self-checking calorific-value signal than either alone, with each able to flag when the other drifts. The estimator succeeds despite sparse instrumentation; with better instrumentation it would not be discarded but strengthened.

Live deployment. Chapter 8 already established the embedded form of the estimator: a laboratory Raspberry Pi prototype with a web dashboard, event detection, and confidence intervals, all tested in shadow mode and integrated with the digital-model workflow. The remaining work is the industrial validation of that already demonstrated architecture: migration to the target embedded platform, live read-only OPC-UA acquisition from the historian, and long-term operation on real plant data.

From advisory to decision support. Once deployed, the estimator would begin as an advisory signal: it presents an estimate, and the operator decides. A natural next stage is to record, alongside each estimate, the action the operator actually took, building an evidence base for how well the estimate's guidance holds up in practice. Only once that record shows the estimate to be reliably trustworthy would it make sense to let it inform the gas trim more directly, as a feed-forward suggestion the control system or the operator can act on before a fuel-quality swing disturbs the temperature, rather than after. The wording matters: the estimate suggests, it does not control. That is the sequence industrial practice trusts, and it is the realistic path from a soft sensor toward anticipatory control.

A reusable methodology. The framework of this thesis is not, at root, specific to calorific value. It is a method for recovering an unmeasured quantity from the transient response a disturbance provokes in a controlled process. The same approach, audit the observability, build a calibrated twin to supply the missing labelled events, recover the signal from the disturbance shape, could be turned toward other quantities a cement plant cares about and cannot easily measure: combustion efficiency, fuel burnout, kiln heat demand, or broader indicators of process health. The twin, similarly, is more than a data generator. Once calibrated, it is a sandbox in which the effect of a changed fuel recipe, a different substitution rate, or a shifted operating point can be examined offline, before it is tried on the plant.

Toward a generalisable estimator. The most ambitious horizon is to move from one kiln to many. The calibration here is specific to Oggaz, but the method is not. With twins calibrated to several plants in a group, and a shared feature representation across them, a more general estimator could be learned that transfers to a new kiln with limited site-specific tuning. That is a genuine research direction in its own right, and it points from a single industrial result toward a broadly applicable framework for alternative-fuel monitoring across cement production.

Closing

What the plant gains from this work is not yet a sensor it can switch on. It is something more specific and, for now, more honest: the knowledge that the fuel-quality signal it has been missing is genuinely present in the data it already collects, that it lives in the transients rather than the averages, that it can be read from the shape of those transients independently of the feed-mass uncertainty that has defeated simpler attempts, and that it can be read early enough in the day to be acted on. The principal obstacle was never the absence of recoverable information; it was the availability of suitable labelled data. Chapter 8 then shows that this recovery path can already be carried onto embedded hardware in a validated laboratory prototype, and the remaining step is live industrial validation on the plant. As further operating campaigns are accumulated and new measurements become available through continued collaboration with the industrial partner, the methodology presented here can grow from an offline demonstrator into an operational decision-support tool. In that evolution lies the real value of the work: not as a finished product, but as the foundation of a practical soft-sensing framework for alternative-fuel monitoring in industrial cement production. The harder question, whether the quantity was there to be recovered at all, has been answered.

Bibliography

- [1] A. C. Kahawalage, M. C. Melaaen, and L.-A. Tokheim, “Opportunities and challenges of using SRF as an alternative fuel in the cement industry,” *Detritus*, 2023, version of record January 2023.
- [2] M. C. Samolada and A. A. Zabaniotou, “Energetic valorization of SRF in dedicated plants and cement kilns and guidelines for application in Greece and Cyprus,” *Waste Management*, vol. 34, no. 2, pp. 411–420, 2014.
- [3] R. Sarc, K. E. Lorber, R. Pomberger, M. Rogetzer, and E. M. Sipple, “Design, quality, and quality assurance of solid recovered fuels for the substitution of fossil feedstock in the cement industry,” *Waste Management & Research*, vol. 32, no. 7, pp. 565–585, 2014.
- [4] L. Fortuna, S. Graziani, A. Rizzo, and M. G. Xibilia, *Soft Sensors for Monitoring and Control of Industrial Processes*, ser. Advances in Industrial Control. London: Springer, 2007.
- [5] Y. Jiang, S. Yin, J. Dong, and O. Kaynak, “A review on soft sensors for monitoring, control, and optimization of industrial processes,” *IEEE Sensors Journal*, vol. 21, no. 11, pp. 12 868–12 881, 2021.
- [6] Y. Liu, M. Jia, D. Xu, T. Yang, and Y. Yao, “Physics-guided graph learning soft sensor for chemical processes,” *Chemometrics and Intelligent Laboratory Systems*, 2024.
- [7] M. M. Mateus, J. M. Bordado, and R. G. d. Santos, “Simplified multiple linear regression models for the estimation of heating values of refuse derived fuels,” *Fuel*, vol. 294, p. 120541, 2021.
- [8] J. Tahir, R. Ahmad, and Z. Tian, “Calorific value prediction models of processed refuse derived fuel using ultimate analysis,” *Biofuels*, vol. 14, no. 1, pp. 69–78, 2023.
- [9] M. Lim, Y. Lee, Y. Lee, W. Yang, and S. Kim, “Improving waste-incineration energy recovery efficiency using a reverse calculation algorithm to estimate waste composition and heating value,” *Waste Management*, 2024.
- [10] H. Mikulčić, E. von Berg, M. Vujanović, and N. Duić, “Numerical study of co-firing pulverized coal and biomass inside a cement calciner,” *Waste Management & Research*, vol. 34, no. 8, pp. 866–877, 2016.

- [11] J. H. Friedman, “Greedy function approximation: A gradient boosting machine,” *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
- [12] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD)*, 2016, pp. 785–794.
- [13] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, “LightGBM: A highly efficient gradient boosting decision tree,” in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017, pp. 3146–3154.
- [14] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [15] R. Fatahi, H. Nasiri, E. Dadfar, and S. C. Chelgani, “Modeling of energy consumption factors for an industrial cement vertical roller mill by SHAP-XGBoost: a “conscious lab” approach,” *Scientific Reports*, vol. 12, p. 7543, 2022.
- [16] F. J. Bargagli-Stoffi, G. Cevolani, and G. Gnecco, “Simple models in complex worlds: Occam’s razor and statistical learning theory,” *Minds and Machines*, vol. 32, pp. 13–42, 2022.
- [17] M. Dubova, S. Chandramouli, G. Gigerenzer, P. Grünwald, W. Holmes, T. Lombrozo, M. Marelli, S. Musslick, B. Nicenboim, L. N. Ross, R. Shiffrin, M. White, E.-J. Wagenmakers, P.-C. Bürkner, and S. J. Sloman, “Is Ockham’s razor losing its edge? New perspectives on the principle of model parsimony,” *Proceedings of the National Academy of Sciences*, 2025.
- [18] T. G. Dietterich, R. H. Lathrop, and T. Lozano-Pérez, “Solving the multiple instance problem with axis-parallel rectangles,” *Artificial Intelligence*, vol. 89, no. 1–2, pp. 31–71, 1997.
- [19] X. Chen, P. Qiu, W. Zhu, H. Li, H. Wang, A. Sotiras, Y. Wang, and A. Razi, “TimeMIL: Advancing multivariate time series classification via a time-aware multiple instance learning,” in *Proc. 41st Int. Conf. on Machine Learning (ICML)*, ser. PMLR, vol. 235, 2024, pp. 7190–7206.
- [20] W. Kritzing, M. Karner, G. Traar, J. Henjes, and W. Sihn, “Digital twin in manufacturing: A categorical literature review and classification,” *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018.
- [21] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, “Characterising the digital twin: A systematic literature review,” *CIRP Journal of Manufacturing Science and Technology*, vol. 29, pp. 36–52, 2020.
- [22] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, “Digital twin in industry: State-of-the-art,” *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415, 2019.

- [23] Y. Guo and B. Huang, “State estimation incorporating infrequent, delayed and integral measurements,” *Automatica*, vol. 58, pp. 32–38, 2015.
- [24] R. Dürr, S. Duvigneau, C. Seidel, A. Kienle, and A. Bück, “Multi-rate data fusion for state and parameter estimation in (Bio-)Chemical process engineering,” *Processes*, vol. 9, no. 11, p. 1990, 2021.
- [25] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1999.
- [26] I. Gustavsson, L. Ljung, and T. Söderström, “Identification of processes in closed loop — identifiability and accuracy aspects,” *Automatica*, vol. 13, no. 1, pp. 59–75, 1977, often cited as Gustavsson, Ljung & Söderström 1974; the *Automatica* paper is 1977. VERIFY which you intend to cite.
- [27] C. Bergmeir and J. M. Benítez, “On the use of cross-validation for time series predictor evaluation,” *Information Sciences*, vol. 191, pp. 192–213, 2012.
- [28] V. Cerqueira, L. Torgo, and I. Mozetič, “Evaluating time series forecasting models: an empirical study on performance estimation methods,” *Machine Learning*, vol. 109, no. 11, pp. 1997–2028, 2020.
- [29] D. R. Roberts, V. Bahn, S. Ciuti, M. S. Boyce, J. Elith, G. Guillera-Arroita, S. Hauenstein, J. J. Lahoz-Monfort, B. Schröder, W. Thuiller, D. I. Warton, B. A. Wintle, F. Hartig, and C. F. Dormann, “Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure,” *Ecography*, vol. 40, no. 8, pp. 913–929, 2017.
- [30] OPC Foundation, “OPC unified architecture specification, part 1: Overview and concepts,” Standard, oPC Foundation.
- [31] Raspberry Pi Ltd., “Raspberry pi compute module 4 datasheet,” Datasheet, industrial variant.
- [32] V. Vovk, A. Gammerman, and G. Shafer, *Algorithmic Learning in a Random World*. New York: Springer, 2005.
- [33] A. N. Angelopoulos and S. Bates, “Conformal prediction: A gentle introduction,” *Foundations and Trends in Machine Learning*, vol. 16, no. 4, pp. 494–591, 2023.
- [34] International Electrotechnical Commission, “IEC 60812: Failure modes and effects analysis (FMEA and FMECA),” Standard, 2018.
- [35] —, “IEC 62443-3-3: Industrial communication networks — network and system security — system security requirements and security levels,” Standard, 2013.
- [36] Raspberry Pi Ltd., “Raspberry pi 4 model b datasheet,” Datasheet.