Beyond Measurements

Interpreting Hidden Environmental Patterns with Machine Learning

Katarzyna Szramowiat-Sala, Dr Eng.
Department of Fuel Technology, Faculty of Energy and Fuels
AGH University of Krakow, Poland

Measurements are increasing — understanding is not automatic

The central challenge is to convert process and environmental data into physically meaningful information.

CONVENTIONAL INTERPRETATION

direct readings
thresholds
descriptive statistics

→

WHAT IS OFTEN HIDDEN

nonlinear relationships

temporal dependencies

latent structures

uncertainty & variability

→

ML-ASSISTED INTERPRETATION

pattern-oriented environmental analytics

Scientific question: What can machine learning reveal that is not directly visible in raw measurements?

From prediction difficulty to a scientific question

To illustrate the idea of **“going beyond measurements”**, I use organic gaseous compounds **(OGC)** emissions from small-scale combustion as a focused case study.

01 PREVIOUS EVIDENCE

OGC – difficult to predict

Earlier temporal deep-learning results for an automatic pellet-fuelled boiler

NN-LSTM	$R^2 = -0.51$
CNN-LSTM	$R^2 = 0.01$

J. Cleaner Prod. 472 (2024) 143473
ECM:X 28 (2025) 101322

02 WHY IS OGC DIFFICULT?

INCOMPLETE COMBUSTION

↓

devolatilisation · oxidation
thermochemical reactions

↓

many organic species

↓

OGC
aggregated organic signal

03 THIS STUDY

Instead of asking only:

Can we predict OGC concentration?

I progressively ask:

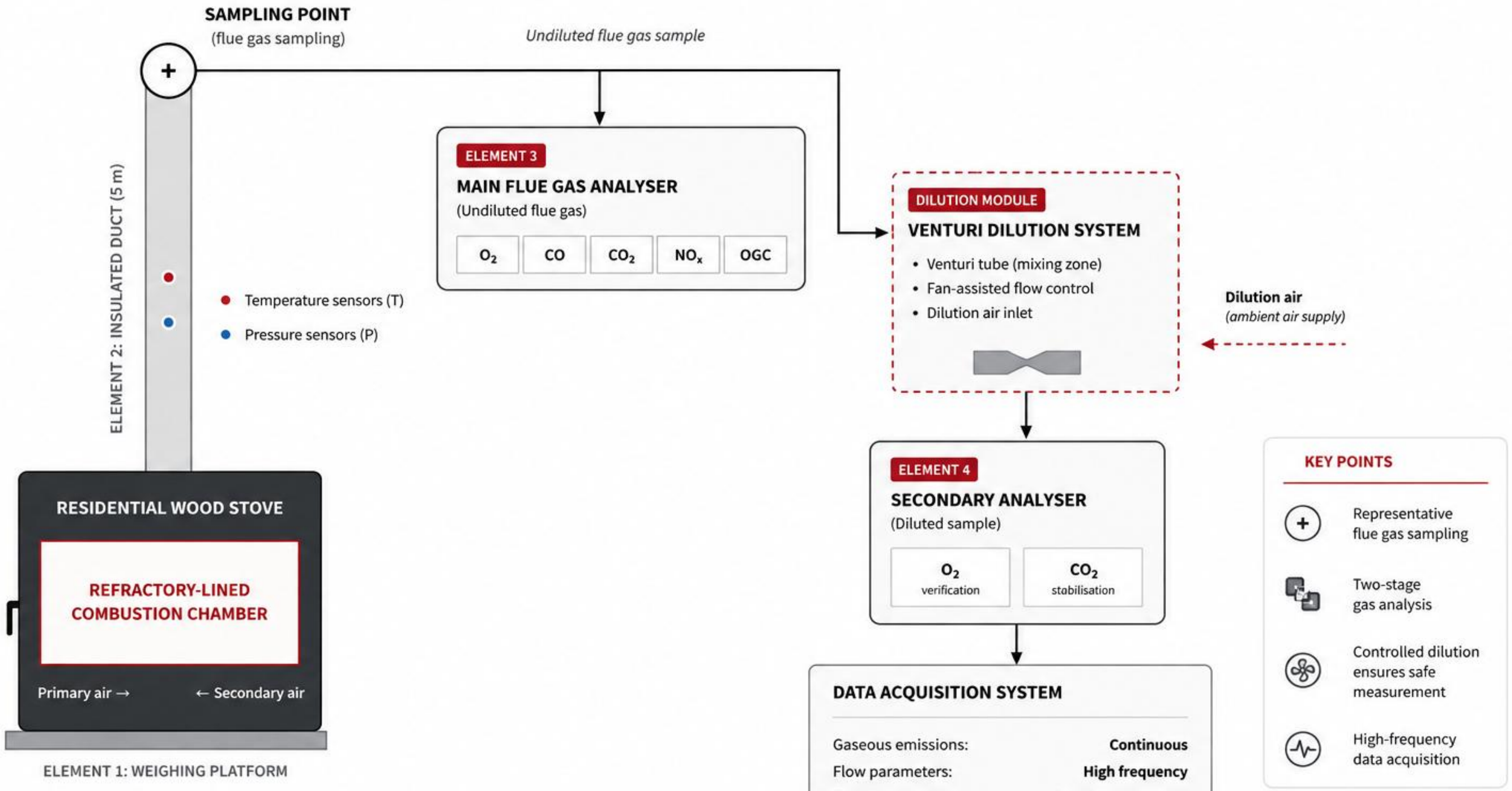
1. **How much OGC?**
2. **Is a high-emission event occurring?**
3. **Can we detect its precursor?**

SCIENTIFIC IDEA

The difficulty of OGC prediction is not only a modelling problem – it is a reason to examine the temporal process information hidden behind the measured concentration.

Experimental Setup & Measurement Protocol

Operational and emission performance of the woodstove was continously monitored.



The system enables accurate measurement of emissions and process parameters under controlled, repeatable conditions.

From Experiment to ML Dataset

STEP 01	STEP 02	STEP 03	STEP 04	STEP 05
Combustion Experiment	Process Measurements	Time Alignment	Quality Control	ML-Ready Series

5 INDEPENDENT EXPERIMENTS

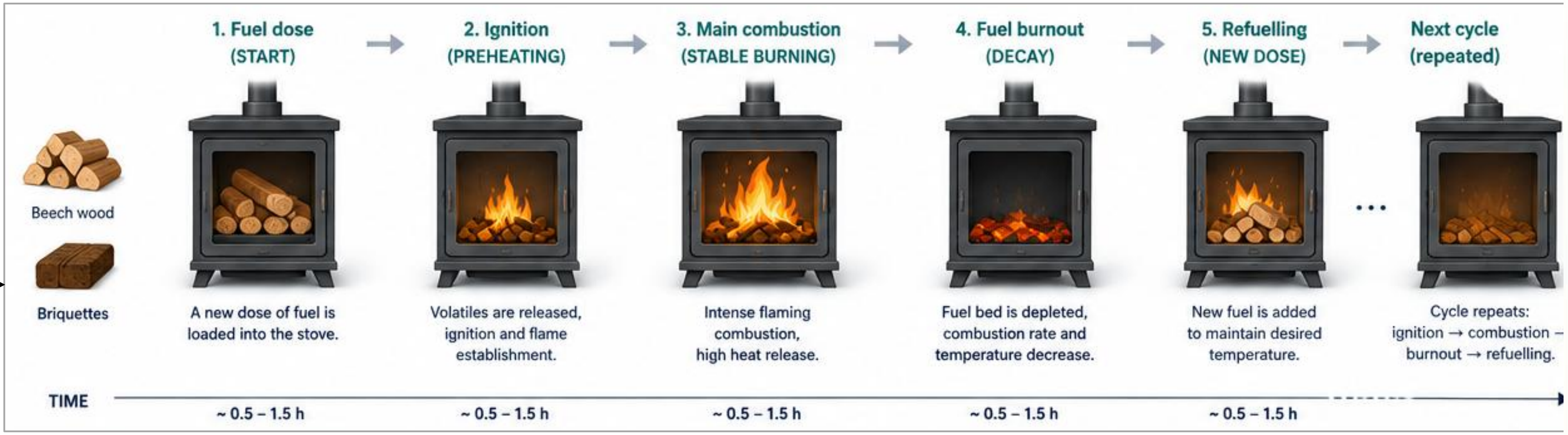
BW1 BW2 BW3

briq2 briq3

Varying wood batch types (beech wood vs briquettes) and stove operational profiles.

TIME-INDEXED MEASUREMENTS FRAGMENT

Timestamp	T_flue [°C]	O_2 [%]	CO [ppm]	CO_2 [%]	Draught [Pa]	OGC [mg/Nm³]
12:01:00	245.2	11.4	842	8.9	14.2	42.1
12:02:00	245.8	11.2	915	9.1	14.1	58.4
12:03:00	246.3	10.9	1120	9.3	14.0	



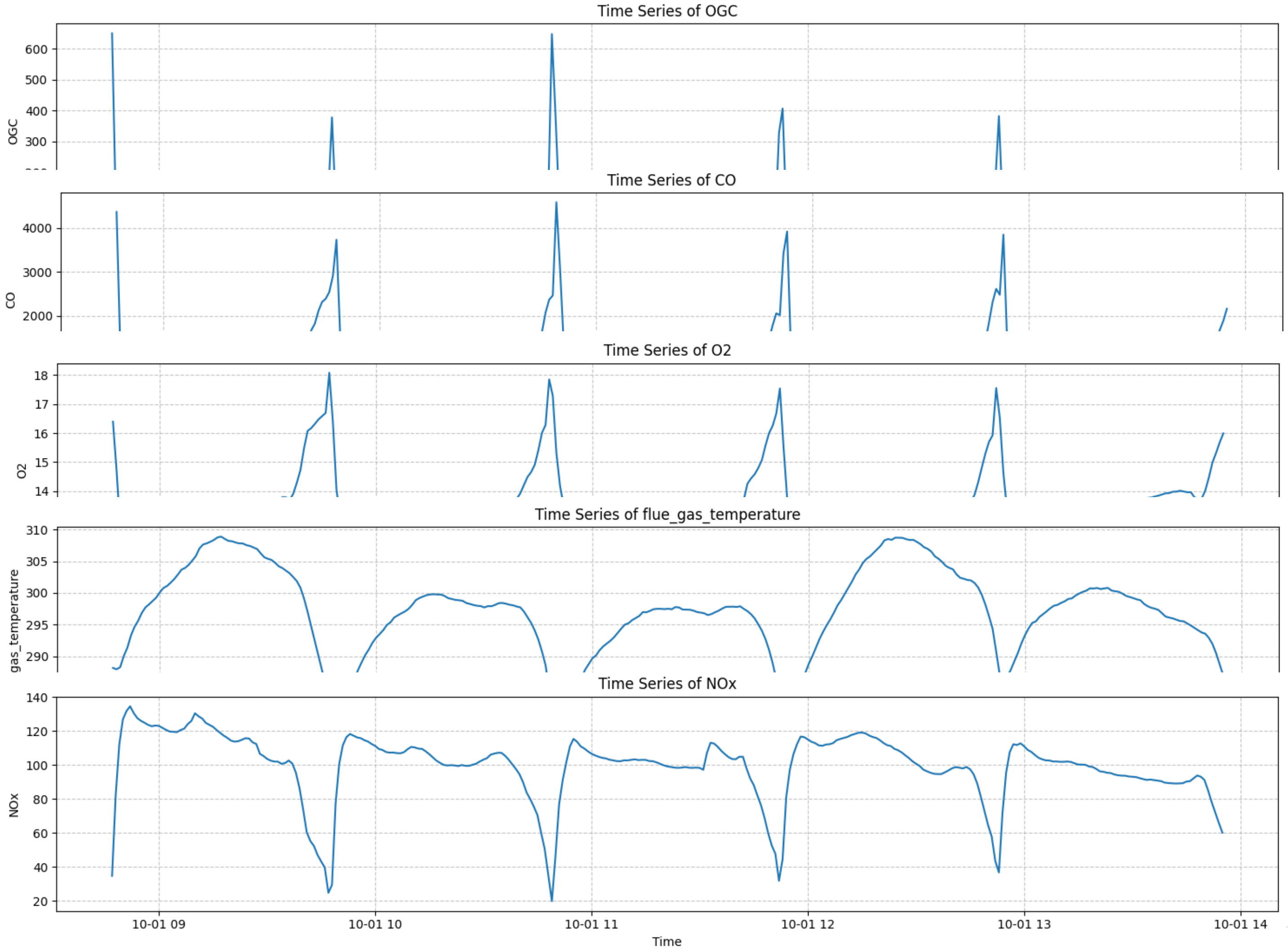
WHY REFUELLING IS NECESSARY?

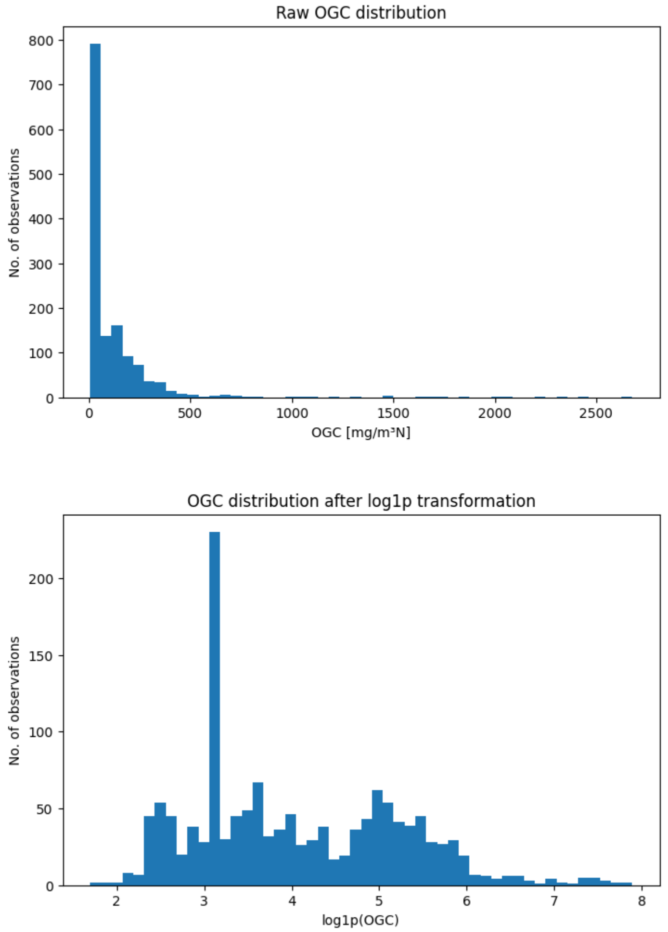
- To maintain desired indoor temperature
- To sustain sufficient heat release
- To avoid excessive temperature drop

Refuelling creates a cyclic and dynamic combustion process

OGC distribution is highly skewed and episodic

Rare high values are not random noise; they form temporally persistent emission episodes.





Raw and log-transformed OGC distributions

OGC dataset: short experiments, sparse extreme peaks

Five combustion tests show high temporal resolution, but strongly uneven emission behaviour.

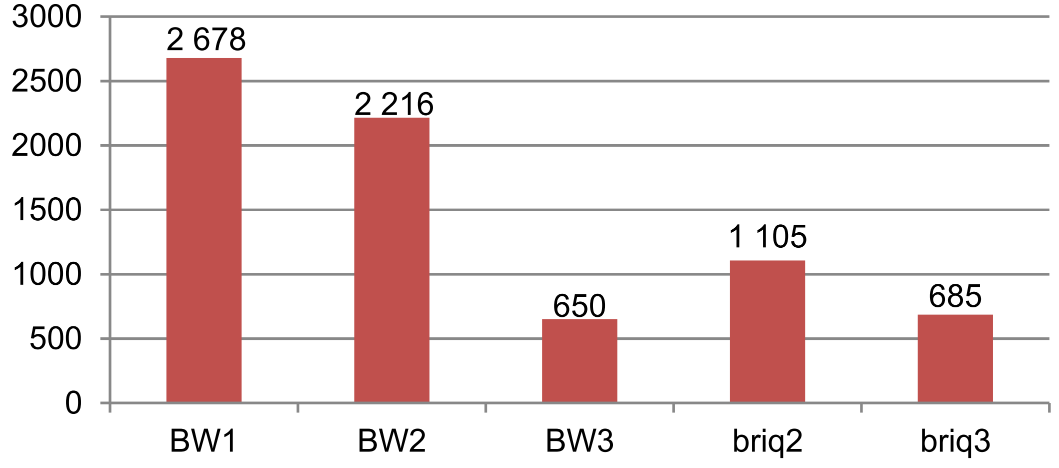
5 combustion tests

1395 1-min observations

2 fuel groups

2678 max OGC [mg m^{-3}N]

OGC maxima by test



Observation from data audit

Beech-wood tests BW1–BW2 contained the highest OGC peaks.
Median values remained much lower than maxima, indicating a heavy-tailed signal.
The analytical target is therefore not a smooth concentration field but a transient emission process.

A small number of high-emission minutes can dominate environmental interpretation.

To Make the Dataset ML-Ready...

...and representative for dynamic and temporal process

...and model with better generalization capabilities

RAW SIGNALS
time-stamped process and emission measurements

REPRESENTATION
state + dynamics + history

SPLIT
Chronological by experiment

VALIDATION
LOTO = Leave-One-Test-Out:
4 experiments → unseen 5th

70% TRAIN

15% VALIDATION

15% TEST

Experiment	n	Iteration	Training set	Held-out test
BW1	n=262	ITERATION 1	BW2, BW3, briq2, briq3	BW1
BW2	n=306	ITERATION 2	BW1, BW3, briq2, briq3	BW2
BW3	n=309	ITERATION 3	BW1, BW2, briq2, briq3	BW3
briq2	n=254	ITERATION 4	BW1, BW2, BW3, briq3	briq2
briq3	n=264	ITERATION 5	BW1, BW2, BW3, briq2	briq3

ambient/context

TRAIN

VALIDATION

TEST

The modelling question evolved with the evidence

The target changed because the engineering meaning of the prediction changed.

01

REGRESSION

How much OGC?

Ridge → RF → XGBoost → LSTM

Baseline reproduced; extreme peaks remained difficult.

Better metric ≠ stronger evidence

The tempting result: longer LSTM window

Window	n test	R^2	MAE	RMSE
5 min	185	0.006	78.6	170.6
10 min	160	0.187	60.4	110.5
20 min	110	0.433	46.4	64.8
30 min	**60**	**0.736**	**28.9**	**42.5**

MAE and RMSE in mg/m³N.

$R^2 = 0.736$ **but nTEST = 60**

Why I did not select the 30-min LSTM:

- only 60 test sequences remained
- long overlapping windows increase redundancy
- **LOTO validation performance was poor ($R^2 = -0.076$)**
- our target was transfer to a new combustion experiment

MODELLING STEP / MODEL | FEATURE REPRESENTATION

01 **Baseline regression**
Ridge regression
current process state
T, O_2, draught, ambient / context

Can static measurements explain OGC?

02 **Temporal feature engineering**
Ridge → Random Forest
state + dynamics + history
ΔT, ΔO_2 • lags • rolling mean / SD

A snapshot is not enough.

03 **Nonlinear regression**
Random Forest → XGBoost
X_full: 18 features
process state + temporal representation

Baseline reproduced; extreme peaks missed.

04 **Sequential modelling**
LSTM
5 / 10 / 20 / 30 min windows
raw OGC • Huber loss • LSTM(32)

Longer context tested; generalization remained key.

CO dynamics SIGNIFICANT in OGC modelling

From LSTM to XGBoost again: Adding physically meaningful CO dynamics improved both average performance and robustness across unseen combustion experiments.

05

Cross-experiment generalization

XGBoost + LOTO

CO-enhanced: 23 features

baseline + CO, CO lags, ΔCO, rolling CO

Train on 4 experiments → test on unseen 5th.

0.339

mean LOTO R^2 · 18 features

→

0.663

mean LOTO R^2 · 23 features

156.2 → 106.8

mean RMSE [mg/m^3N]

ρ = 0.835

global CO–OGC Spearman

XGBoost Performance

Test set	**0.785**	**49.8**	**136.4**

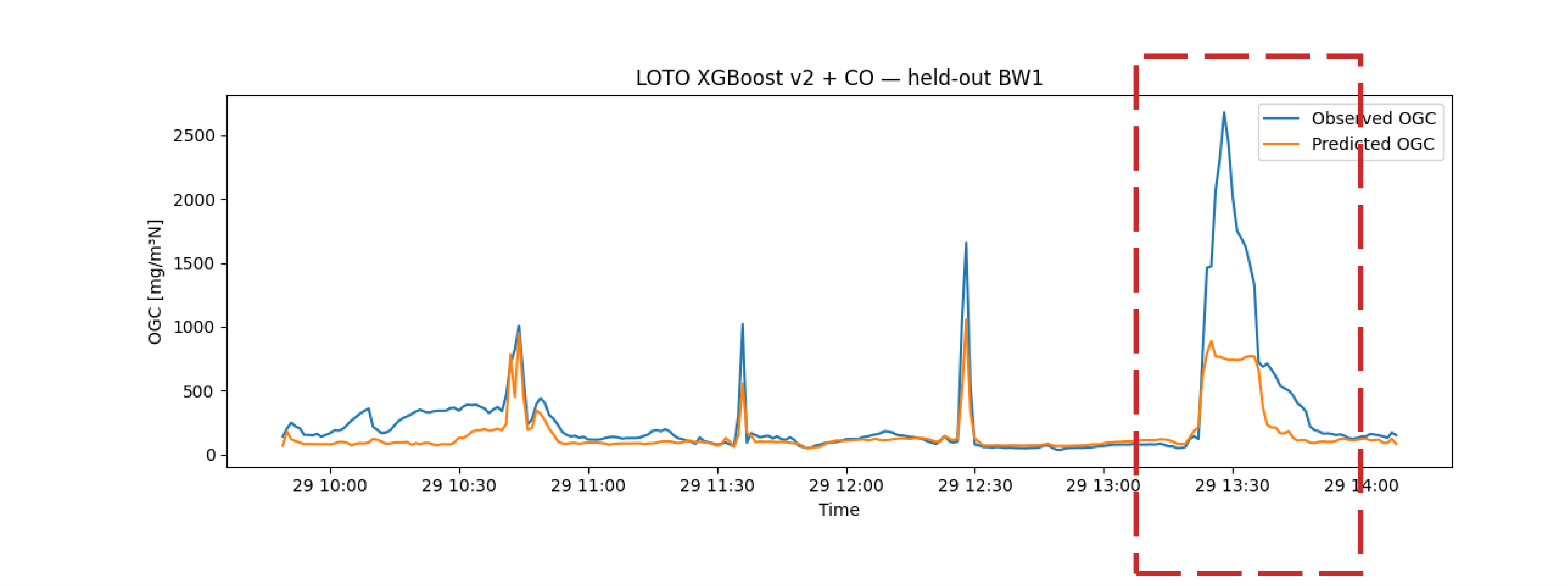
Leave-One-Test-Out (LOTO)

BW1 BW2 BW3 BR2 **BR3** **HELD OUT**

Feature representation	Mean R^2	MAE	RMSE
Baseline	0.339	82.3	156.2
+ CO	0.635	55.4	114.8
+ CO + lag_1	0.650	52.9	110.6
Full CO dynamics	**0.663**	**51.2**	**106.8**

Regression improved – but exact peak height remained difficult

This motivated reformulating the task from point-wise concentration to event detection.



The modelling question evolved with the evidence

The target changed because the engineering meaning of the prediction changed.

01

REGRESSION

How much OGC?

Ridge → RF → XGBoost → LSTM

Baseline reproduced; extreme peaks remained difficult.

02

EVENT DETECTION

Is a critical episode occurring?

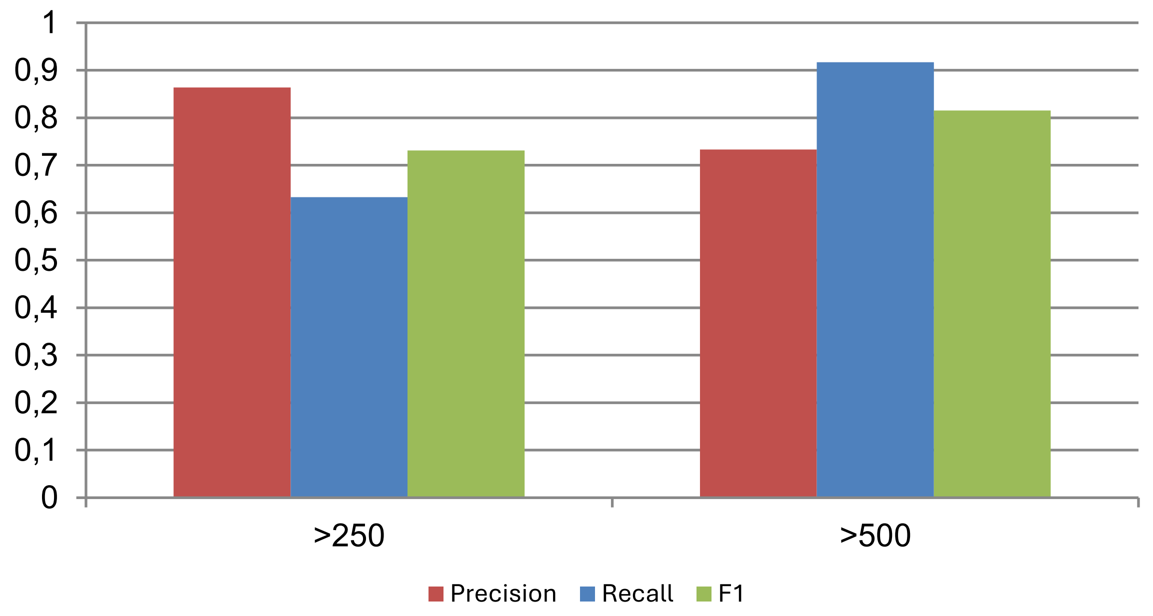
OGC > 250 / 500 mg/m^{3}N

Evaluate point-wise and episode-level detection.

Episode-level evaluation captures environmental relevance

For rare emission states, episode detection can be more meaningful than exact pointwise reproduction.

Episode-level metrics under LOTO



11 / 12

observed >500 mg/m^{3}N episodes detected

0.917

episode recall for >500 mg/m^{3}N

0.815

episode F1 for >500 mg/m^{3}N

06 **High-emission event detection**

Threshold-based classification

OGC > 250 / 500 mg/m^{3}N

point-wise + episode-level evaluation

Exact peak height → detect critical episode.

Key result: regression magnitude was imperfect, but episode recognition was strong enough to support process monitoring.

The modelling question evolved with the evidence

The target changed because the engineering meaning of the prediction changed.

01

REGRESSION

How much OGC?

Ridge → RF → XGBoost → LSTM

Baseline reproduced; extreme peaks remained difficult.

02

EVENT DETECTION

Is a critical episode occurring?

OGC > 250 / 500 mg/m^{3}N

Evaluate point-wise and episode-level detection.

03

EARLY WARNING

Will OGC increase soon?

ΔOGC ≥ 50 mg/m^{3}N; t+1...t+5

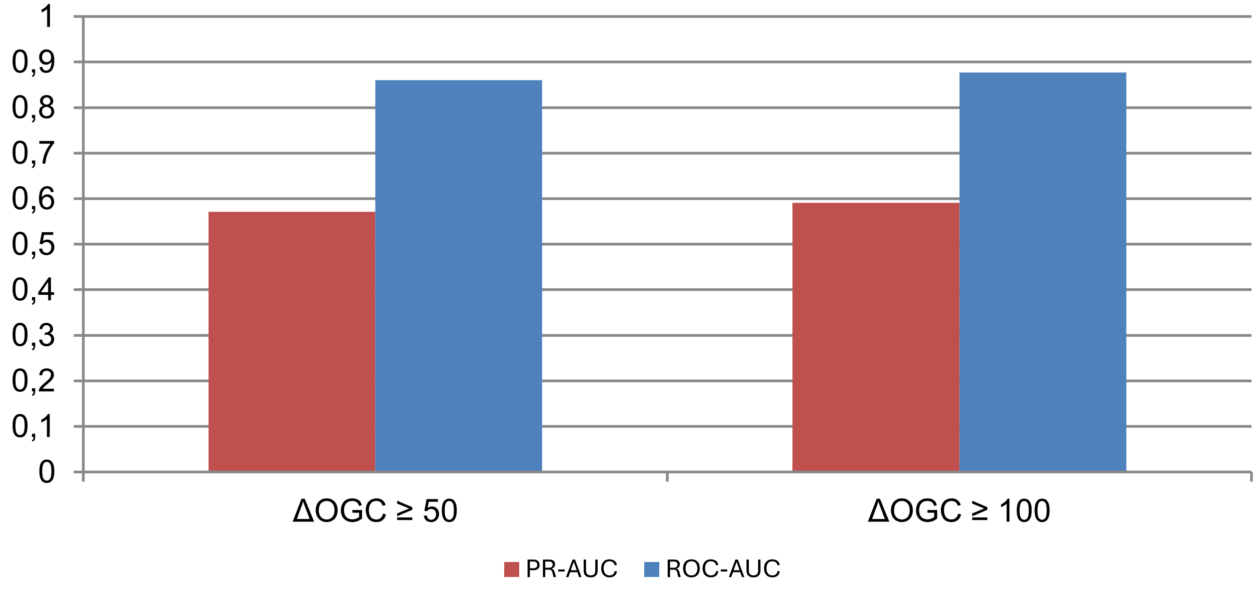
Shift from reproducing a peak to anticipating a hazardous transition.

From detection to early warning

The final engineering question is whether a future OGC increase can be anticipated before it occurs.

Predict whether ΔOGC ≥ 50 or 100 mg m^{-3}N at t + 1 min

Early-warning classification performance



07

Prospective early warning

XGBClassifier + LOTO

predict ΔOGC ≥ 50 mg/m³N

t+1, t+2, t+3, t+5 min • no OGC leakage

Can we warn before the emission surge?

HORIZON	PR-AUC	ROC-AUC
t + 1 min	**0.571**	**0.860**
t + 2 min	0.356	0.787
t + 3 min	0.318	0.812
t + 5 min	0.242	0.758

From concentration prediction to process intelligence

What did the deep data-driven analysis of the combustion process actually add?

01 PROCESS DYNAMICS MATTER
Instantaneous measurements do not capture the full information carried by a transient combustion process. Dynamics and process history are essential parts of the signal.

02 PHYSICAL CONTEXT IMPROVES INTERPRETATION
Combining OGC with physically related process variables — particularly CO dynamics — exposes repeatable relationships between combustion state and organic emissions.

03 EVENTS CAN BE MORE INFORMATIVE THAN PEAK HEIGHT
For highly variable emissions, identifying a critical transition or abnormal episode can be more robust and operationally meaningful than reproducing the exact concentration maximum.

04 GENERALISATION MUST BE TESTED ACROSS EXPERIMENTS
Independent-test validation distinguishes transferable process patterns from relationships specific to a single combustion trajectory.

WHERE THE DATA POINT NEXT

DEEP PROCESS DATA ANALYSIS — **TRANSIENT PATTERNS & PRECURSORS** — **ANOMALY / EVENT DETECTION** — **EARLY WARNING**

The fireplace is a challenging experimental testbed; the transferable outcome is a data-driven framework for detecting abnormal process states and anticipating critical events.

THANK YOU FOR YOUR ATTENTION

Katarzyna SZRAMOWIAT-SALA
Katarzyna.Szramowiat@agh.edu.pl

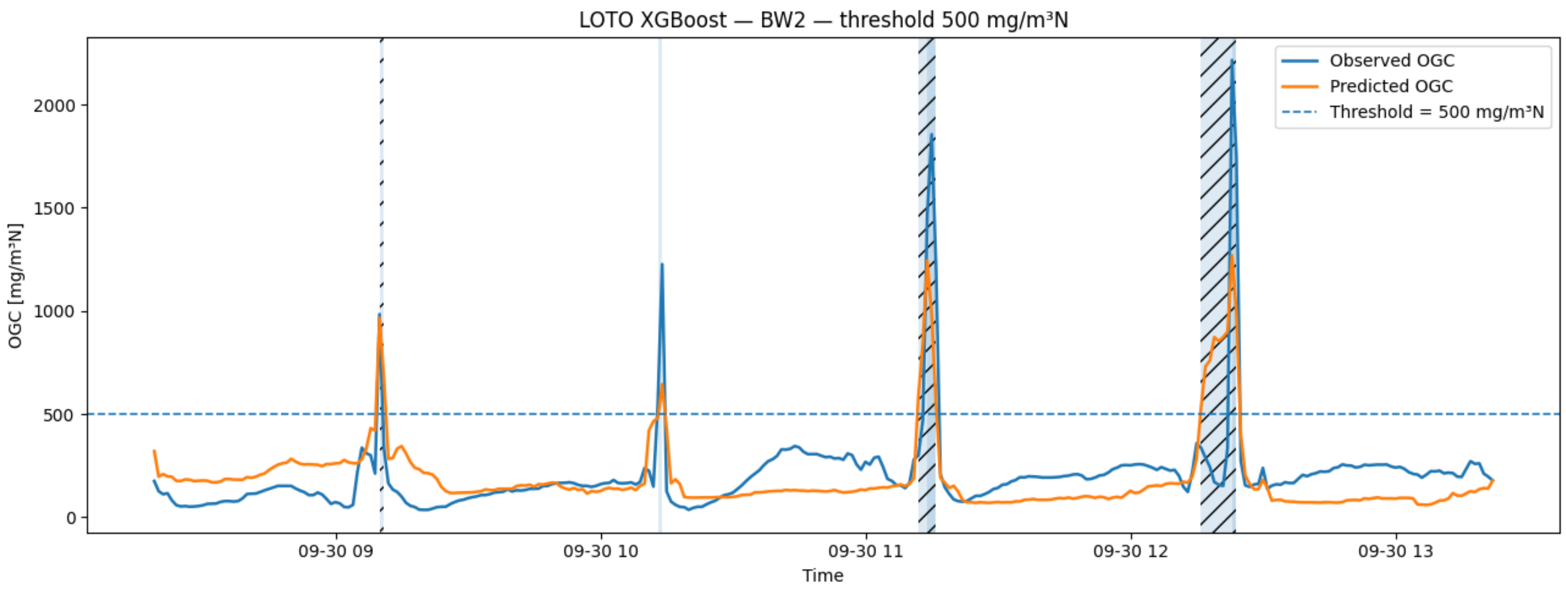
Beyond Measurements

Interpreting Hidden Environmental Patterns with Machine Learning

A Data-Driven and Machine-Learning-Based Assessment of Transient OGC Emission Predictability during Residential Wood Combustion: From Soft Sensing to Early Warning
Katarzyna Szramowiat-Sala, Kamil Krpec, Jiří Ryšavý, Jerzy Górecki

Reframing OGC as a high-emission event signal

The model may miss exact peak height but still detect the environmentally important episode.



Example: LOTO prediction for an unseen test

Interpretation rule

Thresholds
250 / 500 mg m^{-3}N

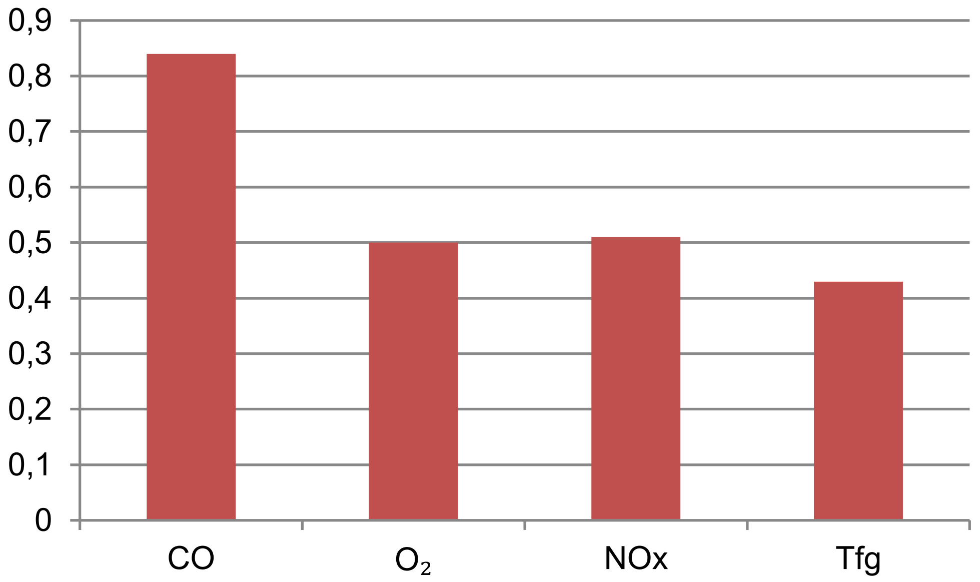
Observed and predicted episodes are compared, not only single minutes

Temporal tolerance accounts for small shifts in process response

Temporal context changes OGC interpretation

Lag analysis indicates that process variables carry delayed or state-dependent relations to OGC.

Best absolute lagged correlations with OGC



Lag behaviour across tests

CO
lag 0 min
|r| ≈ 0.78–0.96

O_2
lag 0–2 min
|r| up to 0.62

NOx
lag 1–10 min
negative relation

Tfg
lag 0–9 min
negative relation

Implication
feature engineering must encode history, not only current values

CO is treated as a predictive marker of the combustion state, not as a causal driver of OGC.

My Features Development Process

Stage 1 — Input Data

Original Merged Dataset 16 columns

PROCESS MEASUREMENTS 6

T_{flue} | O_2 | draught | P_{amb} | T_{lab} | RH

EMISSION MEASUREMENTS 5

CO | NOx | SO_2 | CO_2 | @GC = TARGET

CONTEXT / METADATA 3

fuel | date/time | test_ids

Not all available measurements were used as model inputs.

PRESENT IN DATASET BUT EXCLUDED FROM FEATURES:

~~NOx~~ | ~~SO_2~~ | ~~CO_2~~

Feature engineering ≠ using every available variable

Stage 2 — Temporal Representation

X_full (Baseline) 18 features

CURRENT STATE 8

T_{flue} | O_2 | draught | P_{amb} | T_{lab} | RH | fuel | $t_{elapsed}$

DYNAMICS 2

ΔT | ΔO_2

1^{st} order time derivatives

PROCESS HISTORY 4

T(t-1) | T(t-2) | O_2(t-1) | O_2(t-2)

ROLLING STATS 4

T_{mean} | T_{sd} | $O_{2\,mean}$ | $O_{2\,sd}$

Lags & Rolling Windows encode memory of recent process history

8 Basic Process + 10 Temporal Features = 18 Features

Stage 3 — Enhanced Set

X_full_v2 (CO-Enhanced) 23 features

RETAINED BASELINE FEATURES 18

✓ All 18 features from X_full fully preserved

ADDED CO-RELATED FEATURES +5

+ CO(t) | + CO(t-1) | + CO(t-2) | + ΔCO | + CO_{roll_mean}

Scientific Context: CO and its temporal metrics serve as predictive indicators of combustion instability dynamics (non-causal).

18 Base Features + 5 CO Dynamic Features = 23 Features

STATE + MEMORY + DYNAMICS

The model was not only told **what the process state was** — it was also told **how the process was changing**

Feature engineering transformed raw measurements into a temporal representation of the combustion process.

How the OGC Modelling Strategy Evolved

REGRESSION → EVENTS → EARLY WARNING

	MODELLING STEP / MODEL	FEATURE REPRESENTATION	WHAT WE LEARNED / WHY THE NEXT STEP?
01	**Baseline regression** Ridge regression	**current process state** T, O_2, draught, ambient / context	Can static measurements explain OGC?
02	**Temporal feature engineering** Ridge → Random Forest	**state + dynamics + history** ΔT, ΔO_2 • lags • rolling mean / SD	A snapshot is not enough.
03	**Nonlinear regression** Random Forest → XGBoost	**X_full: 18 features** process state + temporal representation	Baseline reproduced; extreme peaks missed.
04	**Sequential modelling** LSTM	**5 / 10 / 20 / 30 min windows** raw OGC • Huber loss • LSTM(32)	Longer context tested; generalization remained key.
05	**Cross-experiment generalization** XGBoost + LOTO	**CO-enhanced: 23 features** baseline + CO, CO lags, ΔCO, rolling CO	**Train on 4 experiments → test on unseen 5th.**
06	**High-emission event detection** Threshold-based classification	**OGC > 250 / 500 mg/m³N** point-wise + episode-level evaluation	**Exact peak height → detect critical episode.**
07	**Prospective early warning** XGBClassifier + LOTO	**predict ΔOGC ≥ 50 mg/m³N** t+1, t+2, t+3, t+5 min • no OGC leakage	**Can we warn before the emission surge?**

The modelling question evolved: How much OGC? → Is a critical event occurring? → Will an emission surge occur soon?