

Politecnico di Torino

Politecnico di Torino

Master Degree in Electronic Engineering

A.a. 2025/2026

Graduation Session October 2026

Feasibility of Wireless Low-Cost Tracer-Gas Sensor Networks for Multi-Zone Ventilation Assessment

Supervisor:
Prof. Vincenzo Maria Gentile

Co-Supervisor:
Prof. Marco Perino

Candidate:
Raja Daud Ur Rehman Yaqoob

Abstract

Ventilation governs indoor contaminant concentrations, the energy required to condition incoming air, and the risk of airborne transmission. It is measured either by fan pressurisation of the envelope at an imposed pressure difference, or by tracer-gas techniques applied while the building operates normally. Tracer-gas methodology has been settled since the mid-1990s, but its instrumentation is expensive, which confines a measurement to one instrument in one place for a short period. Inexpensive sensors have separately made dense, continuous, distributed measurement feasible. This thesis examines whether low-cost sensing hardware can support tracer-gas measurement with sufficient accuracy and reliability to complement or replace reference instrumentation.

The review follows a documented protocol in which ten AI-assisted search platforms were scored against an eight-parameter rubric before any was adopted, and no tool decided inclusion. A six-branch keyword taxonomy was quantified against the retrieved literature rather than asserted, and every record was scored on four criteria, two from metadata and two from full-text indicators. The search identified 229 records, of which 226 form the analysis corpus and 173 the synthesis corpus.

The publication activity is bimodal, concentrating between 1980 and 1995 and again after 2015, and the two peaks are largely separate literatures. Most consequentially, within the 226-record analysis corpus 98 records address tracer-gas methods and 74 address low-cost sensing, but only nine address both, three of those name the detection principle, and two report a comparison against a reference instrument. Citation chasing added 19 records and no further work, which indicates a property of the field rather than of the search.

Six low-cost tracer platforms were examined in detail. Analytical feasibility is demonstrated for three pathways, namely infrared detection of sulfur hexafluoride within $\pm 20\%$ between 1 and $7\,\mathrm{h}^{-1}$, photoionisation detection of isobutene matching a reference near $1\,\mathrm{h}^{-1}$ at a fifth of the cost, and semiconductor detection of a refrigerant tracer with a mean air-change error of 8%. The two strongest reference comparisons were made in a laboratory layout and a 40-litre box, systems with genuine occupied-building records validate themselves largely against internal consistency, and of

five senses of validation in use only reference-instrument agreement tests what a tracer measurement exists to produce.

Systems with occupied-building records validate themselves mostly against internal consistency, and of five senses of validation in use only reference-instrument agreement tests what a tracer measurement exists to produce. Error analysis explains why the limitation is rarely the transducer. Uncertainty in a decay-derived air-change rate is dominated by the tail of the decay, and in the multizone case published analysis reports flow errors of 10 to 651% for concentration errors of 2 to 30%, according to the conditioning of the zone system. Neither result is addressed by a better detector.

The evidence supports a three-layer architecture combining a reference instrument for absolute scale, low-cost nodes for spatial coverage and an inference layer converting concentrations to flows, rather than low-cost networks replacing reference methods. The principal gap is not a further prototype but a controlled field comparison in which a low-cost network operates in an occupied building alongside a reference analyser long enough to expose drift, with concentration error, decay-fit uncertainty, spatial variability and reference agreement reported separately. A phased experimental programme to supply that evidence is specified.

Acknowledgements

I am especially grateful to my supervisor, Vincenzo Maria Gentile , and co-supervisor, Marco Perino, for their support, valuable suggestions, and constructive feedback throughout this work. Their guidance helped me develop a clear and deeper understanding of the thesis topic and improve the quality of this literature.

I would also like to express my appreciation to the faculty members and academic staff of Politecnico di Torino for the knowledge and academic experience I gained during my Master's studies. The concepts and skills developed throughout the program provided an important foundation for carrying out this research.

I am deeply grateful to my family for their continuous encouragement, patience, and support throughout my studies. Their motivation gave me the strength to remain focused and continue working through the challenges of completing my degree.

I would also like to thank my friends and colleagues for their encouragement, discussions, and support during my academic journey.

Finally, I would like to thank everyone who contributed in any way to completing this thesis. I am sincerely grateful for their support and encouragement.

Table of Contents

List of Tables XI

List of Figures XII

List of Equations XIII

List of Acronyms XV

1 Introduction 1
- 1.1 Context . . . 1
- 1.2 Problem Statement . . . 2
- 1.3 Research Questions . . . 2
- 1.4 Scope and Approach . . . 3
- 1.5 Principal Findings . . . 4
- 1.6 Structure of the Thesis . . . 5

2 Background to Ventilation Measurement and Tracer-Gas Theory 7
- 2.1 The Ideal Tracer Gas . . . 7
- 2.2 The Single-Zone Mass Balance . . . 8
  - 2.2.1 Solutions under the three constraints . . . 9
  - 2.2.2 The assumptions, and why they matter more than the detector 10
- 2.3 Test Designs . . . 11
  - 2.3.1 Concentration decay . . . 12
  - 2.3.2 Constant injection . . . 13
  - 2.3.3 Constant concentration . . . 13
- 2.4 Criteria for Tracer-Gas Selection . . . 13
  - 2.4.1 Physical and transport properties . . . 14
  - 2.4.2 Background concentration and detectability . . . 14
  - 2.4.3 Health, safety and environmental burden . . . 14
  - 2.4.4 Practical use in buildings . . . 15
- 2.5 Blower-Door and Tracer-Gas Testing as Different Physical States . . . 15

3 Methodology for Paper Selection and Scoring 17
3.1 Research Questions . . . 17
3.2 Two Linked Scoring Instruments . . . 18
3.3 Literature Search Strategy . . . 19
3.3.1 Databases . . . 19
3.3.2 Search Strings . . . 19
3.3.3 Discovery and Retrieval Strategy . . . 21
3.3.4 Discovery layer (AI-assisted) . . . 21
3.3.5 Systematic Verification and Retrieval Layer . . . 21
3.4 Inclusion and Exclusion Criteria . . . 22
3.4.1 Inclusion (a record is retained only if all are satisfied) . . . 22
3.4.2 Exclusion (a record is dropped if any apply) . . . 22
3.4.3 Explicit Exception Clauses . . . 22
3.5 Selection Procedure (PRISMA) . . . 23
3.6 Source-Selection Score . . . 26
3.6.1 Generic Score Bands . . . 26
3.6.2 Selection Criteria . . . 27
3.6.3 Threshold and Use of Score . . . 28
3.6.4 Application of the Rubric to the Corpus . . . 28
3.7 Scoring Rubric for the AI Literature-Survey Tools . . . 31
3.8 Survey of AI Literature-Search Tools . . . 33
3.8.1 Scored Comparison and Tools Adopted . . . 34
3.9 Keyword Taxonomy . . . 36
3.10 Limitations of the Search Strategy . . . 38

4 Results of the Literature Search and Corpus Analysis 39
4.1 Composition of the Corpus . . . 39
4.1.1 Temporal Distribution . . . 40
4.1.2 Publication Venues . . . 42
4.1.3 Citation Impact . . . 43
4.2 Distribution Across the Keyword Taxonomy . . . 44
4.3 The Evidence Funnel . . . 46
4.4 Limitations of the Corpus Analysis . . . 48

5 Tracer Gases for Ventilation Measurement 50
5.1 The Selection Problem . . . 51
5.2 Halogenated Reference Tracers . . . 52
5.2.1 Sulfur hexafluoride . . . 52
5.2.2 Perfluorocarbon tracers . . . 53
5.2.3 What the family shares . . . 54
5.3 Occupant and Combustion Gases . . . 54

5.3.1 Carbon dioxide . . . 55
5.3.2 Nitrous oxide . . . 56
5.3.3 Carbon monoxide . . . 57
5.4 Lower-Burden Alternatives . . . 57
5.4.1 Isobutene and photoionisation detection . . . 57
5.4.2 Refrigerant tracers . . . 57
5.4.3 Light hydrocarbons . . . 58
5.4.4 Noble-gas tracers . . . 59
5.4.5 Passive volatile organic tracers . . . 60
5.5 Comparative Assessment . . . 60
5.5.1 The comparison performed directly . . . 62
5.6 Cross-cutting Sources of Uncertainty . . . 63
5.6.1 Mixing and Sampling Representativeness . . . 63
5.6.2 Time-varying Airflow . . . 63
5.6.3 Source Behaviour . . . 63
5.6.4 Chemistry and Building Materials . . . 63
5.6.5 Instrumentation and Data Reduction . . . 64
5.7 Selection Guidance . . . 64
5.8 Conclusion . . . 65

6 Low-Cost Sensor Technologies for IAQ 66
6.1 Scope and the Maturity Gradient . . . 66
6.1.1 The maturity gradient . . . 67
6.2 Non-Dispersive Infrared Sensing . . . 69
6.2.1 Performance and its limits . . . 69
6.2.2 What carbon dioxide can and cannot represent . . . 70
6.3 Optical Particle Sensing . . . 70
6.4 Metal-Oxide Sensing and the Selectivity Problem . . . 72
6.4.1 Why the output is a proxy . . . 72
6.4.2 Where these sensors are nonetheless useful . . . 73
6.5 Electrochemical and Photoionisation Detection . . . 73
6.6 The System Layer . . . 74
6.7 What the Evidence Supports . . . 74

7 Low-Cost Tracer-Gas Sensing Systems 76
7.1 From Classical Methods to Low-Cost Platforms . . . 76
7.2 The Evidence Boundary . . . 78
7.3 Engineering Comparison of Retrieved Platforms . . . 78
7.4 Semiconductor Sensing of Refrigerant Tracers . . . 80
7.5 Occupant-Generated CO_2 in Real Buildings . . . 81
7.6 Photoionisation Detection of a Deliberate VOC Tracer . . . 83

7.7 NDIR Detection of SF_6 in the TEBE-SENSE Platform 83
7.8 What "Validation" Means Across These Systems 84
7.9 Design Guidance for a Low-Cost Tracer Platform 86
7.10 Technology Selection by Application 89
7.11 Conclusion 90

8 Applications and Diagnostic Use Cases 91
8.1 Four Distinct Measurement Problems 91
8.2 Application Domains 92
8.3 Envelope Leakage and Natural Infiltration 94
8.3.1 Extrapolation from 50 Pa to operating conditions 94
8.3.2 Low-pressure alternatives 94
8.3.3 Defensible comparison requirements 95
8.4 Tall Buildings and Interzonal Leakage 95
8.5 Ductwork and Air-Distribution Leakage 97
8.6 Contaminant-Source Localisation 97
8.6.1 The forward model and why inversion is hard 98
8.6.2 Model-based approaches 98
8.6.3 Sensor placement as part of the inverse problem 99
8.6.4 Apportionment as the achievable form of the problem 99
8.6.5 What low-cost networks can and cannot localise 100
8.7 Method Selection by Diagnostic Goal 100

9 Calibration, Validation and Uncertainty 102
9.1 Uncertainty in the Decay Estimate 102
9.1.1 Fitting the decay 104
9.2 Error Amplification in Multizone Inversion 105
9.3 What Calibration Must Cover 106
9.4 Co-location as the Validation Instrument 107
9.4.1 Separating precision from accuracy 107
9.4.2 Co-location benefits, and what it does not require 108
9.4.3 The size of the correction 108
9.4.4 Laboratory agreement does not predict field agreement . . . 109
9.4.5 Reliability as a measurable quantity 109
9.4.6 Laboratory protocols 109
9.5 A Validation Protocol 110

10 Synthesis and Conclusions 112
10.1 Answers to the Research Questions 112
10.1.1 RQ1. Landscape and maturity 112
10.1.2 RQ2. Difficulties 113

10.1.3 RQ3. Trends . . . 113
10.1.4 RQ4. In-situ challenges . . . 114
10.1.5 RQ5. Tracer selection . . . 114
10.1.6 RQ6. Applications and domains . . . 115
10.2 Strengths the Literature Establishes . . . 115
10.3 Limitations and Research Gaps . . . 116
10.4 The Measurement Architecture the Evidence Supports . . . 117
10.5 The Research Gap . . . 118

11 Future Work and Proposed Experimental Design 119
11.1 Objectives and Rationale . . . 119
11.2 Test Configuration and Phasing . . . 120
11.2.1 Building and zone definition . . . 120
11.2.2 Node classes . . . 120
11.2.3 Phasing . . . 121
11.3 Calibration and Quality Control . . . 122
11.4 Analysis and Reporting Plan . . . 123
11.5 Scope and Feasibility . . . 124

Bibliography 125

List of Tables

2.1 Notation used throughout this thesis 9
2.2 Tracer-gas measurement designs 12

3.1 Scoring Instruments and Their Functions 18
3.2 PRISMA Flow Selection Procedure 24
3.3 Generic paper score bands 27
3.4 The four paper selection criteria 27
3.5 AI Literature-Survey Tools Scoring Rubric 32
3.6 Survey of AI-assisted literature-search platforms 33
3.7 Scored comparison of AI literature-search platforms 35

4.1 Corpus coverage by taxonomy branch 44

5.1 Analytical Comparison of Major Tracer Gases Against Key Evaluation Criteria 61

6.1 Detection principles, targets and dominant failure modes 68

7.1 Pathways to a ventilation estimate, and requirements 77
7.2 Low-cost tracer-sensing platforms in the corpus 79
7.3 Five senses of validation in the low-cost tracer literature 85
7.4 Recommended technology by application, with supporting evidence 89

8.1 The four measurement problems and their reference methods 92
8.2 Application domains and their measurement constraints 93
8.3 Method selection by diagnostic goal 101

9.1 Elements of a validation adequate to a ventilation claim 110

10.1 Three-layer measurement architecture 117

11.1 Reported quantities and proposed acceptance criteria 123

List of Figures

2.1 The single-zone control volume 9
2.2 The three tracer-gas test designs 11

3.1 PRISMA selection flow 25
3.2 Distribution of source-selection scores 29
3.3 Keyword taxonomy populated against the corpus 37

4.1 Publication activity across the corpus 40
4.2 Principal publication venues 42
4.3 Citation impact against publication year 43
4.4 Keyword-branch co-occurrence 45
4.5 Shift of research emphasis over time 46
4.6 The evidence funnel 47

5.1 The tracer-gas landscape 51

6.1 Measurement maturity against ventilation relevance 69

7.1 The validation ladder 85
7.2 Reference-anchored platform architecture 87

9.1 Propagated uncertainty in the decay-derived air-change rate 104

11.1 Proposed experimental phasing 121

List of Equations

2.1 Single-zone tracer mass balance 8
2.2 Mass balance expressed in terms of the air-change rate 8
2.3 Concentration decay solution with the source removed 10
2.4 Steady-state solution under constant injection 10
2.5 Injection rate required to hold concentration constant 10
8.1 Duct volumetric flow from tracer dilution 97
8.2 Forward model relating source strengths to sensor concentrations . . 98
9.1 Two-point estimate of the air-change rate from a decay 103
9.2 First-order propagation of concentration and background error into the air-change rate 103

List of Acronyms

ACH
: Air Change Rate

ACH_{50}
: Air Change Rate at a 50 Pa pressure difference

AI
: Artificial Intelligence

ASTM
: American Society for Testing and Materials

CO
: Carbon Monoxide

CO_2
: Carbon Dioxide

GC
: Gas Chromatography

HCHO
: Formaldehyde

HVAC
: Heating, Ventilation, and Air Conditioning

IAQ
: Indoor Air Quality

IoT

Internet of Things

ISO

International Organization for Standardization

MOS

Metal Oxide Semiconductor

N_2O

Nitrous Oxide

NDIR

Non-Dispersive Infrared

NO_2

Nitrogen Dioxide

O_3

Ozone

PFT

Perfluorocarbon Tracer

PID

Photoionization Detector

PM

Particulate Matter, including PM_1, $PM_{2.5}$ and PM_{10}

PRISMA

Preferred Reporting Items for Systematic Reviews and Meta-Analyses

R134a

Refrigerant 134a (1,1,1,2-Tetrafluoroethane)

RH

Relative Humidity

RQ

Research Question

SF_6

Sulfur Hexafluoride

TVOC

Total Volatile Organic Compounds

VOC

Volatile Organic Compounds

Chapter 1

Introduction

1.1 Context

Ventilation determines the rate at which indoor air is replaced with outdoor air, and therefore governs the concentration of every indoor-generated contaminant, the energy required to condition the incoming air, and the risk of airborne transmission between occupants. Its measurement is consequently a recurring requirement in building physics, exposure assessment, energy auditing and infection control.

Two established methods dominate. Fan pressurisation characterises the building envelope by measuring the flow required to sustain a fixed pressure difference, conventionally 50 Pa, while tracer-gas methods characterise the building in its operating state by following the concentration of a marker gas and inferring the air-change rate from a mass balance. The two are not interchangeable, because they measure different physical states, and converting between them requires an empirical model whose uncertainty is often larger than the measurement uncertainty of either technique, as Section 2.5 sets out.

Tracer-gas measurement is mature, in that the single-zone mass balance, the decay, constant-injection and constant-concentration designs, and the associated measurement standards were all established between roughly 1980 and 1995 and have not required substantial revision since. The limitation of the method has never been its theory but rather the cost of the instrumentation needed to satisfy it, because electron-capture chromatography, photoacoustic spectroscopy, thermal desorption and laboratory gas chromatography are all expensive, and that expense restricts a measurement to one instrument in one location for a short period.

Independently, a second body of work has developed around inexpensive gas and particulate sensors. Falling hardware cost, embedded microcontrollers and wireless communication have made it feasible to deploy tens of sensing nodes where one reference instrument would previously have been placed, and to run them

continuously for months. This literature has grown rapidly since 2015 and now constitutes the majority of published work on indoor air quality measurement.

The combination of the two is the subject of this thesis, which asks whether inexpensive sensing hardware can perform a tracer-gas measurement to a standard that would permit it to substitute for a reference instrument.

1.2 Problem Statement

The proposition is straightforward to state and has not been straightforward to establish. A tracer-gas measurement makes demands that a general-purpose indoor air quality sensor is not designed to meet.

First, it requires selectivity. The instrument must distinguish a deliberately released tracer from the indoor background, and most low-cost gas sensors respond to a class of compounds rather than to a species.

Second, it requires accuracy across a range rather than at a point. A decay experiment traverses the low-concentration end of the sensor's range by construction, and the exponential fit is most sensitive to error precisely where the signal is smallest.

Third, and most consequentially, it requires that the measured concentration support an inference about airflow. That inference passes through assumptions such as adequate mixing, a known or negligible source term, a stable background and a constant air-change rate over the measurement, none of which the sensor can verify and all of which are frequently violated in occupied buildings, so that errors introduced by assumption failure routinely exceed the detection limit of the instrument [1], [2].

A device that reports indoor concentration accurately has therefore not been shown to measure ventilation. Establishing the second claim requires comparison against a reference instrument, in a real building, over a period long enough to expose drift. This thesis examines how far the published literature has gone toward that demonstration.

1.3 Research Questions

The review is organised around six research questions, which are stated in full in Section 3.1 and summarised below.

1. **RQ1. Landscape and maturity.** Which low-cost sensing technologies and tracer-gas methods are currently used, and how mature is each?
2. **RQ2. Difficulties.** What are the principal technical and non-technical obstacles to their use?
3. **RQ3. Trends.** What directions is current research taking?

4. **RQ4. In-situ challenges.** What difficulties are specific to real buildings as opposed to controlled conditions?
5. **RQ5. Tracer selection.** How do candidate tracer gases compare, and which are viable substitutes for SF_6 and CO_2?
6. **RQ6. Applications and domains.** In which domains and use-cases are these methods deployed, and how do requirements differ?

1.4 Scope and Approach

The work is a systematic literature review conducted according to the protocol in Chapter 3. Its scope is indoor environments; outdoor and ambient air quality is excluded except where a method transfers directly. Both the classical tracer-gas literature and the low-cost sensing literature are within scope, since the object of study is their intersection.

Three features distinguish the method from a conventional narrative review. A note on scope is needed before those features are described. The work began as the development of a low-cost multi-gas tracer instrument and was redirected into a review, and its scope narrowed during execution to the sharper question that the evidence turned out to support, namely whether low-cost sensing hardware can carry out a tracer-gas measurement of ventilation to a standard that would let it complement or replace a reference instrument. Indoor air quality sensing is therefore treated in Chapter 6 as the technology base from which such a measurement would have to be built, rather than as the object of study, and the tracer-gas literature of Chapters 5 and 7 supplies the standard against which that base is judged. The narrowing is a result of the corpus analysis in Chapter 4, because the convergence of the two literatures proved small enough that a broader treatment would have had little evidence to review.

The search was conducted with AI-assisted tools, and those tools were themselves assessed. Ten platforms were trialled against a defined rubric and scored on eight parameters before any was adopted, and no tool was permitted to determine inclusion, since every record proposed was verified in Scopus and retrieved independently. The survey and the scores are reported in Section 3.8.

Selection was scored rather than narrated. Every record was assessed against four criteria on a common scale, with a hard gate on topical relevance and an explicit threshold for inclusion and for depth of treatment. Two criteria are computed from bibliographic metadata and two from the full text, and each score carries the evidence that produced it. The limitations of that instrument, including a case in which it initially misjudged a foundational paper, are reported in Section 3.6.4.

The corpus is characterised quantitatively before its content is discussed. Chapter 4 reports the temporal distribution, venue dispersion, citation profile and keyword coverage of the retrieved literature. These are treated as results in their own right, because the shape of the evidence base constrains what the thematic chapters can claim.

1.5 Principal Findings

The search identified 229 records, of which 173 were available for synthesis with the full text in hand. Three findings from the corpus analysis frame the thesis.

The literature is bimodal. Publication activity concentrates between 1980 and 1995, and again after 2015, with a pronounced trough between. Seventy-one records fall before the primary date window and are retained under an explicit seminal-methodology exception. The two concentrations are not successive phases of one research programme but two largely separate literatures, roughly a generation apart, with different authors, venues and concerns.

The field has no disciplinary home. The 178 records carrying an indexed venue are distributed across 91 outlets, 62 of which contribute a single paper. The largest single contributor accounts for under nine per cent of the corpus. Work on the same physical problem is published to building engineers, exposure scientists, sensor developers and standards committees, and these readerships cite one another only partially.

The intersection under study is very small. Ninety-eight records address tracer-gas methods and 74 address low-cost sensing, but only nine address both, and of those nine only three name the detection principle used while two report a comparison against a reference instrument. This ratio did not change when the corpus was extended by citation chasing, which is the search route most likely to surface work at an intersection, and that indicates the gap is a property of the field rather than an artefact of the search strategy.

The consequence is that low-cost tracer sensing is at a stage where the analytical feasibility of each detection pathway has been demonstrated and the field readiness of none has. Chapter 7 examines the six platforms in the corpus that attempt the combination and finds that the two with the strongest reference comparisons were tested in a laboratory layout and a 40-litre box respectively, while the systems with genuine occupied-building records validate themselves largely against internal consistency.

1.6 Structure of the Thesis

The thesis proceeds in four movements. Chapters 2 and 3 establish the physical and methodological foundations, Chapter 4 characterises the retrieved corpus, Chapters 5 to 7 form the substantive review moving from tracer media through sensing instruments to the systems that combine them, and the remaining chapters address diagnostic applications, validation requirements, the synthesis of gaps, and the experimental programme that would close them. Each chapter is set out below with the research questions it answers.

Chapter 2: Ventilation measurement and tracer-gas theory, covering the mass balance, the three test designs, the four tracer-selection criteria, and the distinction from fan pressurisation. Establishes the physical basis and the assessment framework applied throughout.

Chapter 3: Search strategy, the survey and scoring of AI-assisted tools, the keyword taxonomy, the selection criteria and the PRISMA procedure. Defines the method and makes it reproducible.

Chapter 4: Quantitative characterisation of the retrieved corpus. Establishes what the evidence base can and cannot support.

Chapter 5: Tracer gases assessed individually against the criteria of Section 2.4, with a comparative matrix and the cross-cutting sources of uncertainty. Answers the tracer-selection question (RQ5).

Chapter 6: Low-cost sensing organised by detection principle, with the maturity gradient across targets and the system layer above the transducer. Establishes what the instruments can be trusted to report (RQ1, RQ2).

Chapter 7: The six platforms attempting a low-cost tracer measurement, the evidence boundary between laboratory and occupied-building work, and the five senses of validation. Assesses the intersection directly (RQ3, RQ5).

Chapter 8: The four distinct measurement problems grouped under ventilation measurement, and the diagnostic use cases each method answers. Establishes which method suits which question (RQ6).

Chapter 9: Propagation of concentration and background error through the decay estimate and the multizone inversion, and what a calibration and validation adequate to the claim must contain (RQ2, RQ4).

Chapter 10: Answers to the six research questions, the strengths and limitations the literature establishes, and the measurement architecture the evidence supports.

Chapter 11: A proposed experimental programme specified to supply the field

evidence the corpus lacks.

Chapter 2

Background to Ventilation Measurement and Tracer-Gas Theory

This chapter establishes the physical basis on which the rest of the thesis rests. It sets out what a tracer gas is and what makes one suitable, derives the single-zone mass balance from which every measurement design in the review follows, examines the three principal designs and the conditions each requires, and states the four criteria against which every candidate gas in Chapter 5 is assessed. It closes by separating tracer-gas measurement from fan pressurisation, because the two are routinely compared and measure different physical states.

The organising point of the chapter is that the equations are simple and their assumptions are not. A tracer-gas measurement is an inference rather than a direct reading, because the instrument records a concentration and the ventilation rate is then deduced from it through a model whose validity the instrument cannot verify. Almost every difficulty examined later in this thesis, including imperfect mixing, unstable source rates, background drift and sensor placement, is a failure of one of those assumptions rather than of the detector.

2.1 The Ideal Tracer Gas

Tracer gases are deliberately introduced, released or naturally generated gases whose concentration is followed in order to infer ventilation rate, air exchange, mixing, interzonal transport, contaminant pathways or ventilation effectiveness. The ideal tracer is chemically stable, detectable at low concentration, absent or predictable in background air, unreactive with building materials, safe at working

concentration, and transported exactly as the air parcel under study is transported [3], [4], [5].

No gas satisfies all of these conditions in every building, and this is the central practical fact of the field rather than a minor caveat. Selection is a compromise among physics, available detection technology, cost, safety, environmental burden and the timescale of the measurement [6], [2]. Chapter 5 works through that compromise gas by gas; this chapter establishes the terms on which it is made.

The literature is dominated by five families. Sulfur hexafluoride remains the established short-duration reference tracer; perfluorocarbon tracers are the principal choice for passive, long-term, multizone work; carbon dioxide is the most accessible tracer for occupied buildings and low-cost monitoring; nitrous oxide is a historically important infrared-monitored alternative; and a broader group including isobutene, refrigerants such as R134a, stable krypton and passive volatile organic tracers is explored to reduce cost or environmental burden [7], [8], [2].

2.2 The Single-Zone Mass Balance

Every method in this thesis is a strategy for solving one equation under a different constraint. Consider a zone of volume V treated as a single well-mixed control volume, ventilated at volumetric flow rate Q, containing tracer at concentration $C(t)$, supplied with outdoor air at background concentration C_{out}, and receiving tracer from a source of strength G. Conservation of tracer mass gives

$$V\frac{dC}{dt} = G + Q\,C_{out} - Q\,C(t) \tag{2.1}$$

Figure 2.1 shows the corresponding control volume. It is convenient to work with the air-change rate $\lambda = Q/V$, expressed in air changes per hour, which recasts Equation 2.1 as

$$\frac{dC}{dt} = -\lambda\left(C - C_{out}\right) + \frac{G}{V} \tag{2.2}$$

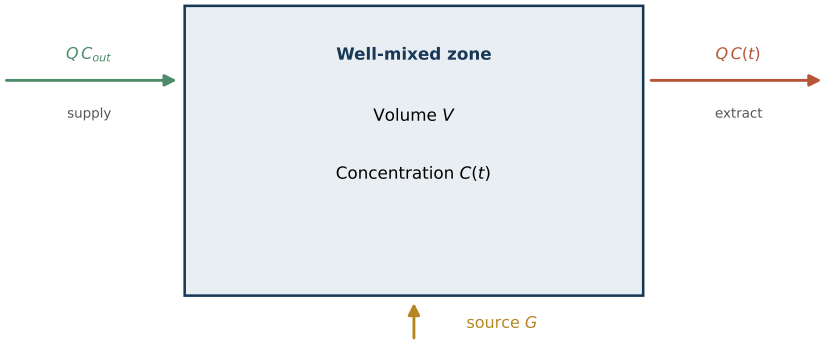


Figure 2.1: The single-zone control volume and the terms of the mass balance of Equation 2.1, in which a zone of volume V is ventilated at flow rate Q, supplied at background concentration C_{out} and fed by a source of strength G. The representation is valid only where the zone is well mixed, where Q is constant over the measurement and where the tracer is conserved within the volume. Every tracer-gas method examined in this thesis solves the same equation under a different constraint, and every source of error examined in Section 5.6 is a violation of one of those three conditions.

Table 2.1 fixes the notation, which is used consistently throughout the thesis. Symbols vary considerably between the sources reviewed here, and the differences are notational rather than physical.

Table 2.1: Notation used throughout this thesis

Symbol	Quantity	Units
V	Zone volume	m^3
Q	Ventilation flow rate	$\mathrm{m}^3\,\mathrm{h}^{-1}$
λ	Air-change rate, $\lambda = Q/V$	h^{-1}
$C(t)$	Indoor tracer concentration	ppm or $\mathrm{mg\,m}^{-3}$
C_0	Indoor concentration at the start of decay	ppm
C_{out}	Outdoor or background concentration	ppm
C_{eq}	Equilibrium concentration under constant release	ppm
G	Tracer source strength	$\mathrm{m}^3\,\mathrm{h}^{-1}$ or $\mathrm{mg\,h}^{-1}$

2.2.1 Solutions under the three constraints

Setting $G = 0$ after an initial release and solving Equation 2.2 gives the decay solution

$$C(t) - C_{out} = \left(C_0 - C_{out}\right) e^{-\lambda t} \tag{2.3}$$

so that a plot of $\ln\left(C - C_{out}\right)$ against time is linear with slope $-\lambda$. Holding G constant instead and allowing the system to reach steady state, where $dC/dt = 0$, gives

$$C_{eq} - C_{out} = \frac{G}{\lambda V} = \frac{G}{Q} \tag{2.4}$$

Holding C constant at a chosen setpoint requires $dC/dt = 0$ with C fixed, so that the injection rate needed at any instant is itself proportional to the ventilation rate,

$$G(t) = Q\left(C_{set} - C_{out}\right) \tag{2.5}$$

which is why this design measures a time-varying ventilation rate directly while the other two return an average.

2.2.2 The assumptions, and why they matter more than the detector

Equations 2.3 to 2.5 are elementary. Their assumptions are not, and four in particular govern whether a measurement is meaningful.

The zone must be adequately represented by the sampling point or network, which is to say that the measured concentration must be the zone's concentration and not a local value. The tracer must not be lost by adsorption onto surfaces or by chemical reaction, or G in Equation 2.1 is not the only source term. The airflow must not change appreciably during the measurement, since λ is treated as constant in the integration. And mixing must be sufficient for the single-zone idealisation to hold at all.

The consequence is stated repeatedly in the methodological literature and is worth stating plainly here, because violations of the mixing, steady-flow or single-zone assumptions routinely produce more error than the nominal detection limit of the instrument [7], [2], [5]. Gao and colleagues quantified one component of this directly, finding that ventilation estimates obtained with different tracer gases can differ by a factor of two to four when dispersion is not passive, which is to say when the tracer's own buoyancy or release momentum contributes to its transport rather than the room airflow alone [1]. An error of that magnitude cannot be corrected by a better detector, because it is not a detection error.

This is the reason the present thesis treats calibration, placement and mixing as parts of the measurement rather than as preliminaries, a position developed in Section 6.7 for sensors and Section 7.9 for complete platforms.

2.3 Test Designs

The three principal designs differ in which term of Equation 2.1 is held fixed and which is measured. Figure 2.2 shows the concentration behaviour each produces, and the contrast explains their very different equipment burdens, since the decay method controls nothing after the initial release whereas holding concentration constant requires the injection rate to track the ventilation rate continuously.

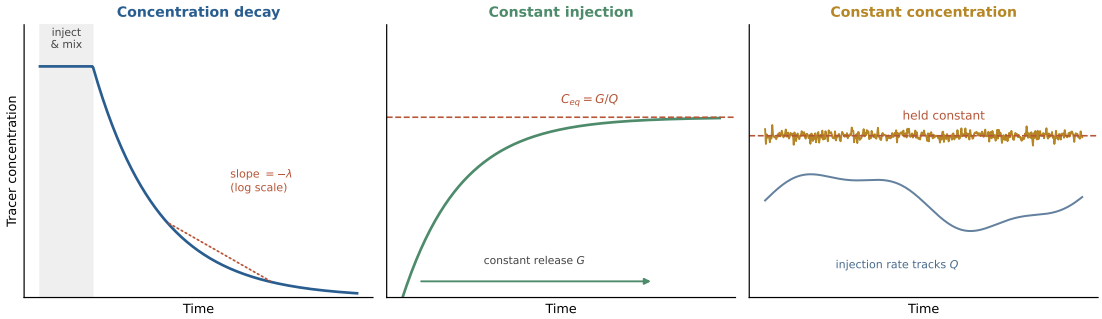


Figure 2.2: Concentration behaviour of the three principal test designs. Decay infers the air-change rate from the exponential slope after a single release, constant injection infers it from the equilibrium concentration $C_{eq} = G/Q$ reached under a sustained known source, and constant concentration infers it from the injection rate required to hold concentration fixed, which is why only the third returns a time-varying rate directly while the other two return an average.

Table 2.2: Tracer-gas measurement designs

Design	Main use	Typical gases	Principal weakness
Concentration decay	Short-term air-change rate and mixing tests	SF_6, CO_2, N_2O, isobutene, CO	Requires adequate initial mixing and a stable decay period [6], [2]
Constant injection	Average ventilation and multizone exchange	PFTs, SF_6, CO_2, N_2O	Requires a known and stable source rate [9], [10]
Constant concentration	Controlled ventilation and large systems	SF_6, CO_2, N_2O	Requires feedback control and continuous analysers [7]
Pulse or step release	Transport time, airflow paths, ventilation effectiveness	SF_6, R134a, CO_2	Interpretation is difficult in complex or reversing flows [11], [12]
Occupant-generated tracer	Realistic occupied-space assessment	CO_2	Source strength and background vary with people and activity [2], [13]

2.3.1 Concentration decay

Decay is the most widely used design because it is the simplest to execute and consumes the least tracer. A quantity of tracer is released, mixed until the zone concentration is acceptably uniform, and the release then stopped, after which the air-change rate follows from the slope of $\ln(C - C_{out})$ against time, as in Equation 2.3. The design is codified in ASTM E741 and ISO 12569, which is a substantial practical advantage because it means a standardised procedure exists against which a measurement can be judged [14], [15].

Its weaknesses follow from what it does not control. The method assumes the zone is well mixed at the moment injection stops, and departures from that condition bias the fitted slope, which Shinohara and colleagues addressed directly by modifying the decay analysis to account for non-uniformity, reporting that the resulting correction reduces error substantially in unevenly mixed spaces [16]. High ventilation rates aggravate the problem, because the decay completes before mixing has been achieved, and large or compartmentalised buildings violate the single-zone premise outright. The method is also sensitive to the background term, since C_{out} enters Equation 2.3 directly and an error in it distorts the whole fit, most severely at the end of the decay, where C approaches C_{out} and the difference between them becomes small. Finally, a single decay gives a single time-specific estimate, and temporal variability in real buildings is large enough that one measurement should

not be treated as characterising a building [2].

2.3.2 Constant injection

Tracer is released at a constant, known rate until removal by ventilation balances the release and the concentration settles at C_{eq}. The air-change rate then follows from Equation 2.4 without any need to resolve a time constant, which makes the design robust to the short-term fluctuations that complicate a decay fit.

Its principal advantage is suitability for long averaging periods and for multizone work, and it is the basis of the passive perfluorocarbon method in which miniature permeation sources supply G and adsorption tubes integrate C over days or weeks [10], [9]. It also links pollutant generation to ventilation naturally, since the mathematics of a tracer source and a pollutant source are identical.

The corresponding requirements are a genuinely stable source rate, since any error in G propagates directly into λ; sufficient time to reach steady state, which rules the design out for short tests; reasonably stable airflow over the averaging period; and a larger total quantity of tracer than a decay test consumes.

2.3.3 Constant concentration

Here the concentration is held at a setpoint C_{set} by continuously modulating the injection rate, and the ventilation rate is read from the injection rate required, following Equation 2.5. Because $G(t)$ tracks Q instantaneously, the design measures a varying ventilation rate as it varies rather than returning an average, which makes it the most informative of the three for buildings whose ventilation is not steady [7].

That capability is bought with equipment. The design requires a continuous analyser, an active feedback controller and a dosing system for each zone, which places it firmly in the research-instrument category and largely explains why it is rare in the low-cost literature reviewed in Chapter 7. It also consumes the most tracer of the three designs, which for a gas with high global warming potential is a substantive objection rather than a matter of cost.

2.4 Criteria for Tracer-Gas Selection

The preceding sections establish how a tracer measurement works, and this section establishes what a tracer must be for that measurement to be valid. The four criteria below are the analytical framework applied to every candidate gas in Chapter 5, and the comparative matrix in Section 5.5 is organised around them. They are stated here rather than in the tracer chapter because they follow directly

from the mass balance, each one corresponding to an assumption that Equation 2.1 makes and that the physical world may decline to honour.

2.4.1 Physical and transport properties

The mass balance assumes the tracer moves with the air. A tracer that disperses under its own buoyancy or the momentum of its release is not following the air parcel, and the measurement then describes the tracer's behaviour rather than the ventilation. Gao and colleagues showed this to be a first-order effect rather than a refinement, with estimates from different tracers diverging by a factor of two to four under non-passive conditions [1].

Passivity is not a fixed property of a gas but a relation between the gas, the release and the room. A smaller density difference from air, a lower release rate and stronger room airflow all push behaviour toward the passive limit; the converse conditions push away from it. A heavy gas released quickly into a quiescent room is the worst case, and it is a case that arises easily in practice.

The tracer must also be chemically stable in the indoor environment, neither reacting with indoor oxidants nor being taken up by surfaces, since either process removes tracer by a route the mass balance does not represent.

2.4.2 Background concentration and detectability

Equation 2.3 depends on the difference $C - C_{out}$, so a tracer with a high or variable indoor background is measured against a moving baseline. This is a decisive criterion in practice rather than a theoretical nicety, since in developing a passive tracer for air-change measurement Paralovo and colleagues considered paraffins and isoparaffins as candidates and rejected them, indoor background being among the properties that disqualified them [17], [18].

The tracer must also be measurable accurately at the concentrations the test will produce. Sulfur hexafluoride owes much of its standing to being detectable at very low concentrations, which permits small releases and short equilibration times. Where the intention is to integrate a measurement over days rather than minutes, compatibility with standard passive samplers becomes a criterion in its own right, because it allows the air-change measurement to share both the timescale and the analytical route of the pollutant measurement it accompanies [17].

2.4.3 Health, safety and environmental burden

A tracer released in an occupied building must be of low toxicity at working concentration and free of known chronic effects, which excludes several otherwise

attractive gases and confines others, such as carbon monoxide, to unoccupied testing.

Environmental burden has become a binding constraint rather than an ethical footnote. Sulfur hexafluoride has an extremely high global warming potential, and the resulting pressure to reduce or replace it is what motivates much of the search for alternatives surveyed in Chapter 5 [2].

Hygroscopic behaviour is a third and less obvious disqualifier. A tracer whose emission rate from a passive source depends on humidity introduces a variable G into a design that assumes a constant one, and candidates have been rejected on precisely this basis during passive-source development [17].

2.4.4 Practical use in buildings

The final criterion is whether a tracer can be used in the building and over the period the study requires. Suitability for occupied spaces and for long campaigns is not implied by good analytical performance, since sulfur hexafluoride is an excellent short-duration reference tracer and yet a poor choice for a month-long measurement in an occupied dwelling, for reasons of environmental burden rather than detection.

Robustness to imperfect mixing belongs here too, because in real buildings mixing is never ideal and the practical question is how badly a given tracer and release configuration degrade when it is not. Source and sampler placement are the levers available, and treating them as part of the measurement design rather than as incidental choices is what distinguishes a defensible field measurement from an optimistic one.

2.5 Blower-Door and Tracer-Gas Testing as Different Physical States

Blower-door testing and tracer-gas testing are frequently compared, and the comparison is only meaningful once it is clear that they measure different things under different conditions.

Blower-door testing pressurises or depressurises a building and measures the airflow required to maintain a specified pressure difference, conventionally 50 Pa, reporting the result as Q_{50}, ACH_{50} or an air permeability. It is a controlled characterisation of the envelope, being reproducible, standardised, and largely independent of weather at the time of the test.

Natural infiltration occurs at pressure differences an order of magnitude smaller and continuously variable, generated by wind, stack effect and mechanical operation. Converting a 50 Pa result into an infiltration rate under operating conditions

therefore requires an empirical leakage–infiltration relationship together with assumptions about climate, shielding, building height and leakiness. The conversion is a model, and its uncertainty belongs to any figure obtained through it.

Tracer-gas testing measures the building in whatever operating state it is in. It captures weather, window position, mechanical ventilation, interzonal exchange and the actual mixing condition at the time of the test, which is its strength. Its corresponding weakness is that the result is time-specific and may change substantially with wind and temperature, which is the same variability that makes a single decay measurement a weak characterisation of a building.

The two methods are therefore complementary rather than interchangeable, and a comparison between them that does not account for the difference in pressure regime and in the model connecting the two is not a comparison of measurement accuracy. Section 4 shows that both traditions are well represented in the corpus, and the studies that compare them directly are examined in the applications material.

Chapter 3

Methodology for Paper Selection and Scoring

This chapter defines how the literature was searched, how individual papers were accepted or rejected, and how both the candidate sources and AI-assisted search tools were scored. The procedure is reproducible so applying the same databases, search strings and criteria should yield a comparable corpus of research. Reporting follows the PRISMA 2020 statement [19], and the appraisal design follows established practice for review-specific quality-assessment instruments [20].

A note on the references cited in this chapter. The works cited here describe the review procedure, namely the reporting standard, the assessment of search systems, the citation-chasing method, the appraisal framework, and the known bias of single-reviewer screening. They are instruments of the method, not objects of the review, and they were never screened against the inclusion criteria of Section 3.4. They are therefore excluded from the corpus export, from the PRISMA flow, from the scoring matrix and from every figure in Chapter 4, and are held in a separate bibliography file so the separation is structural rather than a convention. Every count reported in this thesis, whether the 229 identified or the 173 available for synthesis, refers to reviewed literature alone.

3.1 Research Questions

The review is organised around six research questions, and every results section answers at least one of them, so that the selection and scoring rules below serve a clearly stated purpose rather than collecting papers for their own sake.

1. **RQ1. Landscape and maturity.** Which low-cost sensing technologies and tracer-gas methods are currently used for continuous IAQ and ventilation

monitoring, and how mature is each, pollutant by pollutant?
2. **RQ2. Difficulties.** What are the main technical difficulties (sensing principle, selectivity, drift, calibration) and the non-technical ones (cost, feasibility, logistics, deployment, maintenance)?
3. **RQ3. Trends.** What are the current research trends, for example IoT integration, machine-learning calibration, and emerging low-cost tracer methods?
4. **RQ4. In-situ challenges.** What challenges are specific to in-situ (real-building) assessment as opposed to controlled laboratory conditions, such as imperfect mixing, sensor placement, occupancy and background drift?
5. **RQ5. Tracer selection.** How do candidate tracer gases compare for low-cost continuous ventilation measurement, and which alternative (including hydrocarbon/VOC) tracers are viable substitutes for SF_6 and CO_2, along with the pros and cons of tracer gases and their viability in different conditions?
6. **RQ6. Applications and domains.** In which application domains (homes, schools, offices, smart buildings) and for which use-cases (ventilation diagnostics, leak/infiltration detection, occupancy estimation, exposure assessment) are low-cost IAQ and tracer-based sensing deployed, and how do the requirements differ across them?

3.2 Two Linked Scoring Instruments

The methodology uses one consistent scoring language (a 0–10 scale with explicitly defined bands) for two distinct purposes, and keeping the scale common makes the link between selecting papers and selecting tools transparent.

Table 3.1: Scoring Instruments and Their Functions

Instrument	What it scores → what it decides
Source-selection score	Applied to every candidate paper across four selection criteria 3.6, and it governs inclusion or rejection, and separates papers kept for deep critical analysis from those cited as background.
Tool-evaluation score	Applied to each AI literature-survey tool across the eight parameters already defined 3.7, and it governs which tools were adopted to run the search.

In short, the selection criteria decide which papers enter the corpus and the tool parameters decide which tools were used to find and process them, and both are scored on the same 0–10 bands so the reader sees a single, coherent scoring methodology.

3.3 Literature Search Strategy

3.3.1 Databases

Searches were run on the electronic databases available through Politecnico di Torino with Scopus as the primary index, ScienceDirect (Elsevier) for full-text retrieval, and Google Scholar for forward/backward citation chasing. Scopus was chosen as the primary index because it applies independent editorial content selection and supports the reproducible Boolean querying a stage-by-stage flow requires [21]. Google Scholar is used only for citation chasing and never as the systematic index, since it does not support reproducible Boolean retrieval and is unsuitable as a principal search system for that reason [22].

3.3.2 Search Strings

Boolean strings were built from the keyword taxonomy, and the exact strings together with the date each was run are recorded so that the search can be repeated, with the representative pattern shown below.

```
("low-cost sensor*" OR "low cost sensor*") AND ("indoor air
quality" OR IAQ OR ventilation) AND (CO2 OR "tracer gas" OR VOC
OR "particulate matter" OR NDIR OR "air change rate")
```

1. **S1 Core low-cost IAQ sensing**

```
("low-cost sensor*" OR "low cost sensor*" OR "low-cost monitor*")
AND ("indoor air quality" OR IAQ) AND (monitor* OR measurement
OR sensing)
```

2. **S2 Specific pollutant sensors**

```
(NDIR OR "metal oxide" OR MOS OR electrochemical OR
photoacoustic OR "low-cost sensor*") AND (CO2 OR "carbon
dioxide" OR "particulate matter" OR PM2.5 OR VOC OR "volatile
organic compound*" OR formaldehyde) AND indoor
```

3. **S3 Calibration / validation / uncertainty**

```
("low-cost sensor*" OR "low cost sensor*") AND (calibrat* OR
co-location OR colocation OR validation OR drift OR uncertainty)
AND ("indoor air quality" OR IAQ)
```

4. **S4 IoT / system architecture**

```
("indoor air quality" OR IAQ) AND ("internet of things" OR IoT
OR "sensor network*" OR "real-time monitoring" OR ESP32 OR
Arduino OR "Raspberry Pi")
```

5. **S5 Tracer-gas methods & ventilation rate**

```
("tracer gas" OR "tracer-gas") AND ("air change rate" OR
"ventilation rate" OR ACH OR "air infiltration") AND (decay OR
"constant injection" OR "constant concentration")
```

6. **S6 Specific tracers & alternatives**

```
(SF6 OR "sulfur hexafluoride" OR perfluorocarbon* OR PFT OR
"nitrous oxide" OR N2O OR "CO2 tracer") AND (ventilation OR "air
change rate") AND (building* OR indoor)
```

7. **S7 Application domains**

```
("indoor air quality" OR ventilation) AND (school* OR classroom*
OR office* OR residential OR home* OR "smart building*") AND
(CO2 OR "low-cost sensor*" OR monitor*)
```

8. **S8 Emerging low-cost tracer sensing**

```
("low-cost" OR "low cost" OR wireless OR IoT) AND ("tracer gas"
OR R134a OR "tracer dispersion") AND (ventilation OR "air
change")
```

3.3.3 Discovery and Retrieval Strategy

To adopt innovative tools for the survey, the literature was assembled in two complementary layers. AI-assisted tools were used for discovery, which specifically is finding candidate papers and mapping how they relate, while a conventional bibliographic database provided the systematic basis that makes the search reproducible and supplies the counts reported in the PRISMA flow 3.5. The two layers have distinct strengths where AI tools excel at surfacing relevant and semantically related work quickly, however, they do not expose a fixed, controlled index that another reviewer could re-query identically, and a bibliographic database offers exactly that controlled coverage and reproducibility, but then it gives weaker relevance ranking, so combining them captures the recall of AI discovery with the rigour of a defined index.

3.3.4 Discovery layer (AI-assisted)

AI tools were applied according to the review stage they best support.

- **Discovery and relevance mapping.** Research Rabbit, Connected Papers, Litmaps and Undermind were used to locate candidate papers and to map citation neighbourhoods and research timelines, which also drove the backward/forward citation chasing.
- **Screening and data extraction.** Elicit, Undermind and Consensus were used to read, summarise and extract findings from candidate papers during screening.
- **Citation validation.** Scite was used to check how later work supports or contradicts key papers.

The comparative evaluation of these tools, scored on the eight parameters, is reported in 3.7 and this section concerns only their role in building the corpus.

3.3.5 Systematic Verification and Retrieval Layer

To make the search reproducible and to obtain stage-by-stage counts, every candidate identified through the discovery layer was verified against Scopus as the primary index, and the Boolean strings described in 3.3.2 were run in Scopus to define the systematic record set, to confirm that the AI-discovered papers fall within a defined query, and to identify any indexed work the discovery layer had missed, while full texts were retrieved through ScienceDirect (Elsevier) and the respective publishers via Politecnico di Torino access. The number of records remaining after each step in this layer is what feeds the selection funnel in 3.5.

3.4 Inclusion and Exclusion Criteria

Criteria are written to be operational and each can be checked against a paper with a yes/no answer, with each of them made to stand on their own merits independently of any pre-existing reading list.

3.4.1 Inclusion (a record is retained only if all are satisfied)

- Addresses at least one of indoor air quality measurement, ventilation-rate determination, tracer-gas methods, or low-cost gas/particulate sensing.
- Peer-reviewed journal article or conference paper, or an authoritative review or measurement standard.
- Published within the primary window [2006–2026], with the exception of seminal-works in 3.4.3.
- Reports a method, evaluation, deployment, or review relevant to at least one research question.
- Full text available in English.

3.4.2 Exclusion (a record is dropped if any apply)

- Outdoor or ambient air quality only, with no indoor relevance.
- Purely clinical or health-outcome studies with no sensing or ventilation method.
- No methodological or empirical content (opinion, editorial, marketing material).
- Duplicate records, preprints superseded by a published version, or inaccessible full text.

3.4.3 Explicit Exception Clauses

Seminal-methodology Exception

Foundational tracer-gas and ventilation-measurement papers are retained regardless of publication date when they define a method still in use (for example the single-zone decay, constant-injection and constant-concentration formulations). This is standard practice in systematic reviews and prevents the date window from discarding the methods the field is built on.

Cross-domain Methodology Exception

A study set outside conventional buildings (for example a naturally ventilated agricultural structure) is retained when its tracer-gas methodology transfers directly

to building ventilation, and its eligibility is judged on method and topic relevance, not on the setting.

3.5 Selection Procedure (PRISMA)

Records pass through identification, de-duplication, title/abstract screening, full-text eligibility assessment and final inclusion, and the number remaining at each stage is reported, giving the flow shown in Figure 3.1. The stage names and the requirement to report both the number remaining and the reason for each removal follow the PRISMA 2020 flow [19]. Records added by backward and forward citation chasing are reported as a separate identification route rather than being folded into the database count, following snowballing practice, which treats the chased set as a distinct and separately auditable source [23].

The screening stages that can be decided from bibliographic metadata alone are applied programmatically. A script reads the exported corpus, matches every record against the taxonomy of Section 3.9, evaluates the date window and the seminal-methodology exception of Section 3.4.3, and computes the topical-relevance and currency scores of Section 3.6. It writes one row per record to a screening matrix, with the two criteria that require reading the full text left blank for manual completion. Separating the two makes it visible which decisions were mechanical and which were the reviewer's, and it means the flow diagram regenerates rather than being maintained by hand as the corpus grows.

Populations reported in this thesis. Six counts recur in what follows, and because two of them are cut from the identified set on different criteria rather than in sequence, each is defined here and the same label is used throughout.

Identified (229): every record located, comprising 210 from the database and AI-assisted search and 19 added by citation chasing during retrieval.

Analysis corpus (226): the identified set less the two language exclusions and the one duplicate cite key of Section 3.4. This is the population behind every figure and every count in Chapter 4, including the keyword taxonomy and the evidence funnel, because those measure what the retrieved literature contains rather than the thesis cited.

Screened in (195): the identified set less the 34 records whose topical relevance fell below the hard gate of 5.

Full text retrieved (179): the screened-in set less the 16 records for which no full text was obtained.

Eligibility assessed (177): the retrieved set less the one language exclusion and the one duplicate that were held in full text and removed at this stage.

Synthesis corpus (173): the assessed set less the four records confirmed below the inclusion threshold on reading for deep analysis. This is the population that is cited and carried into the bibliography.

The analysis corpus and the screened-in set are two different cuts of the same 229 records rather than consecutive stages, since the first removes language and duplicate records while the second removes records failing the topical gate, and those two removals do not overlap. The analysis corpus is therefore larger than any stage of the flow below screening, and the scored matrix of Section 3.6.4 contains 212 records, being the identified set less the 17 listed in the retrieval register.

Table 3.2: PRISMA Flow Selection Procedure

Stage	**Records remaining**	**Removed at this stage (reason)**
Identification (database and AI-assisted search)	210	—
Additional records from citation chasing during retrieval	+19	—
Total identified	229	0 (no duplicate cite keys at identification)
Title/abstract screened	195	34 (topical relevance below the hard gate of 5)
Full text retrieved	179	16 (no copy obtainable; two permanently unavailable, one also a language exclusion)
Eligibility assessed	177	2 (one not published in English, one duplicate DOI)
Confirmed below threshold on reading	173	4 (read in full, judged inadequate)
Available for synthesis	**173**	—

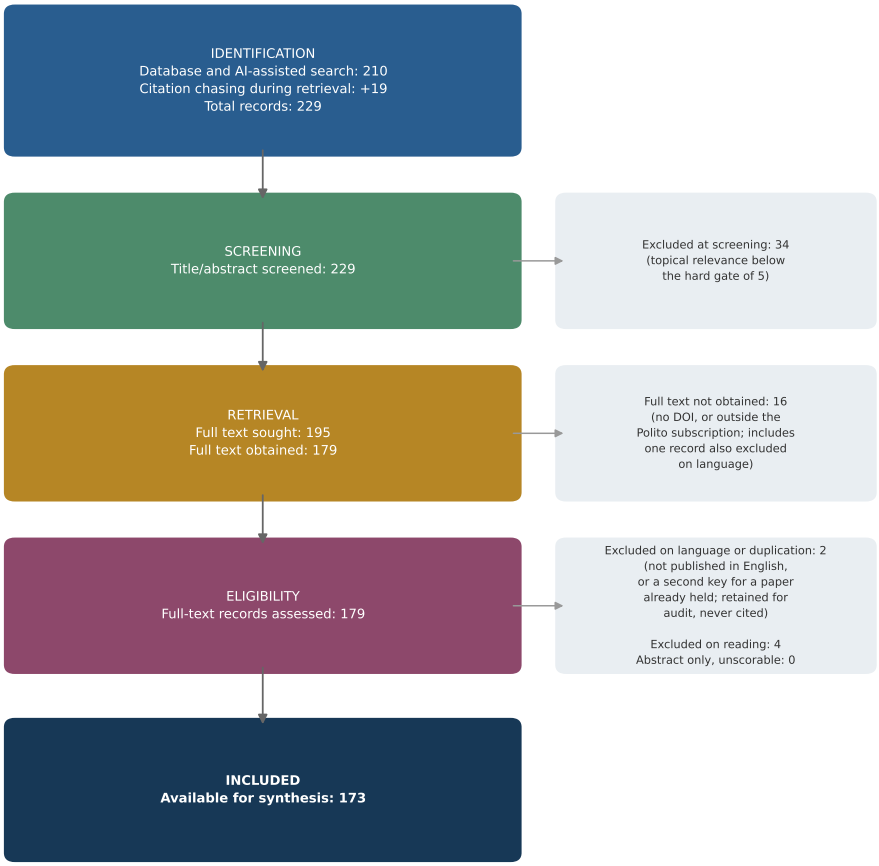


Figure 3.1: PRISMA selection flow for the review, with the number of records remaining at each stage on the left and the reason for each removal on the right. Screening is applied from record metadata, and depth of treatment within the included set follows the full-text scoring of Section 3.6. Every stage is computed from the corpus export and regenerates as the corpus changes. Records excluded on language or duplication are carried through each stage.

Note on completeness. Retrieval is effectively complete, with 179 of the 195 screened records obtained in full text. The 16 that were not are predominantly pre-2000 conference material published without a DOI, mainly in the ASTM, ASHRAE and AIVC series, and they comprise 15 records for which no copy could be traced together with one of the language exclusions, which was never retrieved. Two of the 15 are recorded as permanently unavailable, having proved untraceable through the library, the publisher and the secondhand market alike, and the claims that rested on them were re-sourced to held literature and are documented in the retrieval register rather than silently dropped. The retrieval register itself lists 17 records, being those same 15 together with two whose downloaded file contained

no assessable article, and 17 is therefore the count used in Section 3.6.4, where the scored matrix is defined.

Language and duplicate exclusions. Two records, one in Korean and one in Japanese, satisfy the topical criteria but fail the English full-text requirement of Section 3.4. A third record is a second cite key for a paper already held under a different key, identified by the duplicate-DOI check described in Section 3.5. They leave the flow at two different stages, because one of the two language exclusions was never retrieved and is therefore counted among the records for which no full text was obtained, while the other language exclusion and the duplicate were both held in full text and are removed at the eligibility assessment. They are reported at every stage of Figure 3.1 so that the exclusion is auditable, and they are excluded from the synthesis, from every figure in Chapter 4, and from the bibliography. A check in the analysis script fails the build if either cite key appears anywhere in the thesis source, which prevents them being reintroduced by accident.

Depth of treatment. The full-text scoring of Section 3.6 proposes a depth of treatment within the 173 retained records, separating those subjected to close critical analysis from those cited as supporting or background evidence. That assessment is confirmed or revised as each chapter is written, and the final division is reported in Section 3.6.4.

3.6 Source-Selection Score

Every paper that reaches full-text assessment is scored on four criteria, each on the shared 0–10 scale. The score decides inclusion and the depth of treatment, and it is the formal link between the scoring system and the selection criteria.

3.6.1 Generic Score Bands

The generic score bands given below are the five bands that are used throughout this research on every paper. No standard appraisal instrument exists for a review that spans measurement methodology and sensor hardware, so a review-specific instrument was constructed, which is the recommended course where no validated checklist fits the question; the requirement it must meet is that the criteria and their anchors are stated explicitly enough for another reviewer to apply them [20]. That is the purpose of the bands below and of the per-criterion anchors in Table 3.4.

Table 3.3: Generic paper score bands

Score	Band	Generic meaning
9–10	Excellent	Best-in-class for this criterion; no meaningful gap.
7–8	Good	Clearly satisfies it; minor limitations only.
5–6	Adequate	Usable but with evident limitations.
3–4	Weak	Present but unreliable; heavy reservations.
0–2	Poor/Absent	Largely missing or unusable for this criterion.

3.6.2 Selection Criteria

The four selection criteria given below are used for each paper, described with anchors.

Table 3.4: The four paper selection criteria

Criterion	Low (0–2)	Mid (5–6)	High (9–10)
Topical relevance	Off scope; no link to any RQ.	Touches the topic but addresses an RQ only indirectly.	Directly central to one or more RQs.
Methodological soundness	Method unclear or not described.	Reasonable method with gaps or weak controls.	Rigorous, well-controlled, or a benchmark method.
Evidence / empirical strength	Claim only; little or no data.	Some data but limited scope or validation.	Extensive data and validation against references.
Currency & standing	Dated and superseded (no exception applies).	Older but still cited and relevant.	Recent and influential, or a foundational method (4.3).

3.6.3 Threshold and Use of Score

- **Hard gate.** Topical relevance must be ≥ 5 and a paper below this is excluded regardless of its other scores.
- **Inclusion.** Beyond the gate, a mean score ≥ 6 / 10 across the four criteria includes the paper in the synthesis.
- **Depth of treatment.** A mean score ≥ 8 proposes the record for deep critical analysis, while a mean of 6 - 7.9 proposes it as supporting or background evidence. The threshold proposes rather than decides, since a record may be moved to background use on reading if it duplicates a source already cited, and Section 3.6.4 reports how often that occurred.

3.6.4 Application of the Rubric to the Corpus

The rubric above was applied to every record rather than to a selected subset, and the result is reported as a scored matrix with one row per paper. Two of the four criteria are decided from bibliographic metadata, with topical relevance following from the taxonomy match and the indoor scope, and currency and standing following from the date window, the seminal-methodology exception and the citation rate. The remaining two are scored from a fixed list of textual indicators that a script detects in the retrieved document, such as whether a methods section or an analytical derivation is present, whether a measurement standard is named, whether the work is compared against a reference instrument, and whether error metrics, sample sizes, measurement durations and statistical treatment are reported. The detection is a deterministic pattern match that records whether a marker is present, so it forms no view about the paper and produces a count rather than an assessment. Each score is stored alongside the indicators that produced it, so any individual judgement can be traced back to the paper and checked.

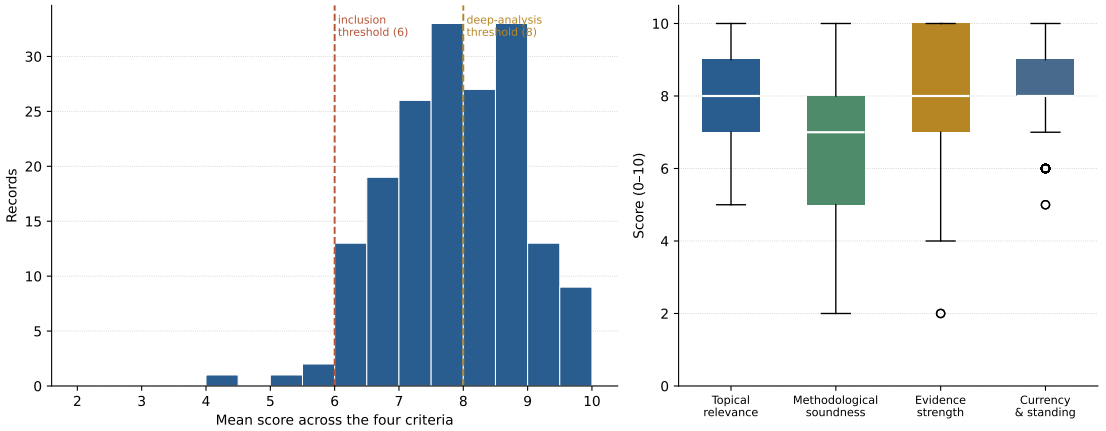


Figure 3.2: Distribution of the mean source-selection score across the four criteria (left) and the spread of each criterion individually (right), for the 177 records scored on the full text.

Applying the thresholds of Section 3.6 to the 177 records scored on the full text gives 82 at or above the deep-analysis threshold of 8, and 91 between 6 and 8 to be cited as supporting or background evidence. The depth actually applied differs from this, and the difference is reported rather than reconciled away. Fourteen records scoring at or above 8 proved on reading to be redundant with a source already cited for the same point, or to concern a subject the thesis does not treat, and were moved to supporting or background use. The final division is therefore 68 records subjected to deep critical analysis and 105 cited as supporting or background evidence. Their scores were left untouched, because the score records what a paper is and it stands, whereas depth of treatment records how this thesis uses it, which is a separate judgement and one the score can inform but not make. Each of the fourteen carries its reason in the scoring matrix. The criterion medians are informative in themselves, since topical relevance, evidence strength and currency all sit at 8, while methodological soundness sits at 7 with a much wider spread, reaching as low as 2. The weakest criterion across the corpus is therefore the description of method, which is consistent with the observation in Section 4.2 that a large share of the literature reports what was measured without specifying how.

What the automated scores are, and are not. The two full-text criteria are reproducible proxies for methodological quality, not a substitute for reading, and they must be understood as such. No automated system read a paper, summarised one, or formed a judgement about its quality at any point in this review, and no record entered or left the corpus on the strength of such a judgement. The indicators rank records and flag them for attention, and the reading and the decisions that follow from it are the author's throughout. An early version of this procedure scored Sherman's *Air Infiltration Measurement Techniques*, a foundational and heavily

cited methodology report, below the inclusion threshold, because the indicators had been written around modern paper structure and did not recognise a 1980s technical report that carries no headed methods section and writes of uncertainty rather than of RMSE. That was a bias in favour of recent typesetting rather than a judgement about quality, and the indicators were rewritten to accept analytical derivation in place of experimental apparatus and sustained error analysis in place of named error statistics.

The episode motivates the rule that governs the matrix, which is that *the automated score never excludes a paper on its own.* Only the topical-relevance hard gate, which is decided from metadata and is reliable, removes records automatically, and a record falling below the inclusion threshold on the full-text indicators is referred for reading instead.

Outcome of the manual review. Twelve records were referred in this way and all twelve were read. Eight were retained with an upward override, and the pattern among them is instructive, since it falls into three groups. Two of them, D'Ottavio's error analysis of perfluorocarbon-derived multizone rates and Sherman's analysis of errors in passive measurement, carry no abstract in the export, so the metadata-derived topical relevance was scored far below what the papers warrant. Three more, including Dietz's Brookhaven multizone perfluorocarbon study, were retrieved as scanned page images from the 1980s and early 1990s, and reading them showed the indicators to have been depressed by the state of the printed page rather than by any weakness in the work. The remaining three, among them Lagus and Persily's review of tracer-gas techniques and a Lawrence Berkeley laboratory evaluation of low-cost particulate monitors against reference instruments, are short documents whose brevity the indicators mistook for thinness. Four were confirmed below the threshold on reading and are recorded as excluded, two of them because the venue proved to have no established editorial record. Each decision carries a written justification in the scoring matrix, so an override can be examined rather than merely asserted.

Eight records in the corpus were retrieved as scanned page images rather than as digital text, and these were read and scored on that reading. They are concentrated in the classical literature, where the printed original is often the only form in which a paper survives. A further two downloads only hold a cover page and abstract, one being a single-page conference poster and the other an abstract-only conference submission, and because there is no article behind either they are treated as unretrieved.

Records excluded for want of a full text. Seventeen records could not be assessed, comprising 15 that passed the topical gate but for which no full text was obtained, and two whose downloaded file contained no assessable article. Methodological soundness and evidence strength cannot be judged without reading the work, so these records are not scored, are not cited anywhere in the thesis, and

are excluded from the bibliography. They remain in the PRISMA flow of Figure 3.1 as unretrieved, which is where such a record belongs, and are listed in full in a separate register so the decision is auditable. The scored matrix therefore contains 212 records, of which 173 are available for citation. That figure is the same 173 reported at the foot of the PRISMA flow, since the diagram and the matrix are generated from one another and cannot disagree.

3.7 Scoring Rubric for the AI Literature-Survey Tools

The eight parameters already presented for the AI tools are scored on the same 0–10 bands (Section 3.6.1), and the anchors below answer the point directly that for each parameter there is a defined meaning at the low, middle and high of the scale, so a score can be reproducible rather than a matter of opinion.

Table 3.5: AI Literature-Survey Tools Scoring Rubric

Parameter	Low (0–2)	Mid (5–6)	High (9–10)
Search quality	Mostly irrelevant hits or broken search.	Finds obvious papers; misses adjacent work or adds noise.	Surfaces nearly all relevant work with little noise.
Papers access	No links to full text.	Some full text; many paywalled gaps.	Reliable full-text links or open-access retrieval.
Quality of summarization	No summary or misleading ones.	Broadly correct but shallow summaries.	Comprehensive, faithful, well-structured summaries.
Evidence extraction	Cannot pull data points or tables.	Extracts simple facts; struggles with tables/figures.	Extracts structured data, tables and findings reliably.
Citation validation	No citation context.	Shows counts only, not supporting/contrasting context.	Distinguishes supporting vs. contrasting citations in context.
Cost effectiveness	Costly with little usable free tier.	Fair value; key features behind a paywall.	Strong value; free tier already useful for the workflow.
Ease of understanding	No help interpreting difficult text.	Some explanation of hard terms.	Clearly explains and simplifies complex passages.
Visualization & usability	Cluttered or unusable interface.	Workable interface; limited mapping/visuals.	Clear interface with useful maps and navigation.

Aggregation. Each tool is reported as a row of eight scores with a total (or weighted total). If some parameters matter more for this thesis like search quality and evidence extraction, the weights for those points are stated explicitly rather than left implicit.

3.8 Survey of AI Literature-Search Tools

Ten platforms were identified and trialled against the rubric in Section 3.7. Each was used on the same seed query drawn from the taxonomy in Section 3.9, so the comparison reflects behaviour on this thesis topic rather than general reputation. Table 3.6 records what each tool costs and what it does well or badly; Table 3.7 converts those observations into the scored form defined above. Pricing is as observed at the time of trial and is quoted per month.

Table 3.6: Survey of AI-assisted literature-search platforms

Platform	Pricing	Strengths	Weaknesses
Research Rabbit	Free tier; $10–12.5	Citation-graph traversal is strong, and the relevance mapping between papers makes backward and forward chasing efficient. Most functions remain usable on the free tier.	Slow in use, with intermittent search faults. Produces no generative summaries, so screening still has to be done by reading.
Semantic Scholar	Free	Broad academic index with frequent open-access full text, and a stable API. Useful as the neutral cross-check against tools with opaque ranking.	No mapping or visualisation of any depth, and no synthesis of retrieved material.
Leapspace	Free trial; institutional licence	Retrieves literature and performs synthesis, and additionally surfaces the researchers and funders active on a topic, which supports collaboration scouting.	Little independent review or validation of its ranking, no transparent pricing, and a small user base.
Elicit	$7 Plus; $29 Pro	Built around systematic-review workflows, with structured extraction into a table of study characteristics.	Free use limited to roughly four searches. Frequently lacks full-text access, so extraction of tables and numerical results often fails.
Consensus	Free tier; $15	Strong at aggregating what a body of papers claims about a specific question. The in-paper search locates the exact supporting sentence, which makes verification quick.	Reports positions rather than interrogating them; little independent critical analysis of methodological quality.

Continued on next page...

Table 3.6: Survey of AI-assisted literature-search platforms (continued)

Platform	Pricing	Strengths	Weaknesses
Scite	$20	The only tool trialled that classifies citations as supporting or contrasting, which is directly useful for judging whether a method has been corroborated or challenged.	Free tier returns only counts, not the citation context that gives the tool its value.
SciSpace	Free tier; $12–20; $70 Advanced	Comprehensive: any sentence can be selected for immediate explanation, which is valuable on unfamiliar analytical chemistry. Useful extraction and drafting functions.	Free credits exhaust quickly, and summary structure is inconsistent between documents.
Litmaps	Free basic; $10	Timeline mapping shows how a topic developed and when methods were superseded, which suited the historical tracer-gas literature.	Essentially a mapping tool; weak at summarisation, extraction and comparison.
Connected Papers	Free (limited); $3–10	Clear graph visualisation with colour encoding for publication date and relevance, which makes clusters and outliers immediately visible.	Discovery only. No analysis, extraction or summarisation.
Undermind	Free tier; $16	Conversational interface that returns pinpoint references rather than paraphrase, and reports its own search coverage. Produced the deepest topic-specific retrieval of the tools trialled.	Less established and less externally validated than the alternatives, with a correspondingly thin record of independent evaluation.

3.8.1 Scored Comparison and Tools Adopted

Applying the eight parameters of Section 3.7 to the platforms above gives Table 3.7. Search quality and evidence extraction are weighted double, because the binding constraint in this review was finding the sparse low-cost tracer-sensing literature and pulling quantitative performance figures out of it, not producing readable prose.

Table 3.7: Scored comparison of AI literature-search platforms

Platform	**SQ**	**PA**	**QS**	**EE**	**CV**	**CE**	**EU**	**VU**	**Wtd.**
Undermind	9	7	9	8	6	8	8	7	**8.1**
SciSpace	7	8	7	8	5	6	9	7	**7.2**
Consensus	7	6	9	6	6	8	8	6	**7.0**
Elicit	7	5	7	8	4	5	7	6	**6.4**
Semantic Scholar	7	8	3	3	5	10	4	4	**5.5**
Scite	5	5	5	5	10	4	5	6	**5.5**
Research Rabbit	7	6	2	2	6	8	3	8	**4.9**
Connected Papers	6	5	2	2	5	8	3	9	**4.6**
Litmaps	5	5	3	2	5	8	3	8	**4.5**
Leapspace	5	5	5	4	3	4	5	5	**4.5**

SQ = search quality, PA = papers access, QS = quality of summarization, EE = evidence extraction, CV = citation validation, CE = cost effectiveness, EU = ease of understanding, VU = visualization and usability. Weighted mean applies a factor of 2 to SQ and EE.

The scores separate the platforms into three groups with different roles, and the review used them accordingly rather than selecting a single winner. Undermind, SciSpace and Consensus form the retrieval and synthesis layer and did the substantive work of locating and interrogating papers. Research Rabbit, Connected Papers and Litmaps form a discovery layer whose value is exclusively in mapping the citation neighbourhood, and their low totals reflect the rubric penalising absent summarisation rather than poor performance at what they actually do. Scite occupies a category of its own, scoring 10 on citation validation and near the bottom elsewhere, and was used only to test whether a contested method had been corroborated or disputed downstream.

Two limitations of this exercise should be stated. The scores are single-rater judgements formed during use on one topic, so they are reproducible against the anchors in Table 3.5 but not independently replicated. The platforms are also under active development, and the pricing and capabilities in Table 3.6 are a snapshot rather than a stable description. More importantly, no tool was permitted to determine inclusion. Every record proposed by an AI platform was verified in Scopus and retrieved independently before being scored under Section 3.6, which keeps the discovery layer separate from the evidence base and is the reason the search strings in Section 3.3.2 are reported in full.

3.9 Keyword Taxonomy

The taxonomy organises the vocabulary of the field into six branches, covering the domain in which a study is set, the ventilation quantity it measures, the tracer it uses, the diagnostic problem it addresses, the sensing technology it employs, and the system-level characteristics of the deployment. It serves two purposes, since it is the source of the Boolean search strings in Section 3.3.2, and it is also the classification scheme under which the retrieved corpus is characterised in Chapter 4.

The tree in Figure 3.3 is populated rather than asserted. An initial vocabulary was drafted from the supervisor's six-branch structure and from reading, then every candidate keyword was matched against the title and abstract of each record in the corpus, and the count of matching records is reported on the leaf. Terms that appeared frequently in the corpus but had no home in the draft tree were added as new leaves; this is how demand-controlled ventilation, machine-learning calibration, blower-door testing and airborne infection control entered the taxonomy. Natural ventilation, absent from the first version of the tree, is now an explicit node under ventilation parameters.



Figure 3.3: Keyword taxonomy for the review, populìated and quantified against the retrieved corpus. Counts on each leaf give the number of corpus records, from the $n = 226$ analysed, whose title or abstract matches that keyword, and a record may appear under several branches, so the branch totals do not sum to the corpus size. The sparsely populated branches are as informative as the well populated ones, and they're interpreted in Section 4.2.

Reporting the tree with counts attached makes it diagnostic rather than decorative. A leaf carrying a high count marks a well-populated part of the field; a leaf carrying a low one marks either a genuine research gap or a defect in the search vocabulary, and the two can be distinguished by inspecting the records concerned. The interpretation of the resulting distribution, including the branches that turn out to be thinly populated, is given in Section 4.2.

3.10 Limitations of the Search Strategy

- English-only inclusion may omit relevant non-English studies.
- Reliance on indexed databases under-represents grey literature and some standards bodies.
- Keyword-driven retrieval can miss papers that use non-standard terminology for the same concepts, however, citation chasing partly mitigates this.
- The seminal-works exception introduces a degree of reviewer judgement, which is documented case by case for transparency.
- Screening and scoring were performed by a single reviewer rather than by two independent reviewers. Single screening is known to miss a proportion of relevant records that double screening retains, and the size of that loss varies widely between reviews [24]. The programmatic screening of Section 3.5 reduces the inconsistency of the metadata decisions, since those are applied by rule rather than by hand, and the full audit trail of Section 3.6 records the evidence behind each judged score, but neither substitutes for an independent second reviewer, and the corpus should be read with that limitation in mind.

Chapter 4

Results of the Literature Search and Corpus Analysis

The methodology of Chapter 3 produced a corpus of 229 identified records, of which 226 form the analysis corpus and 173 are available for synthesis after full-text scoring. This chapter characterises that corpus before any of its content is discussed. The purpose is not bibliographic bookkeeping, because the shape of the retrieved literature is itself a result, and three features of it determine what the thematic chapters that follow can and cannot claim, namely a bimodal age distribution, a dispersed set of publication venues, and an intersection between tracer-gas methodology and low-cost sensing that turns out to be very small. Every figure in this chapter is generated directly from the corpus export by the analysis script described in Section 3.5, so the numbers reported here update with the corpus rather than being transcribed by hand.

4.1 Composition of the Corpus

Of the 226 records in the analysis corpus, 172 carry a resolvable DOI and 177 carry a retrievable abstract, and a full-text copy is held for 209. That figure is larger than the 179 retrievals recorded in the PRISMA flow of Figure 3.1 because the flow counts retrieval only for the 195 records that passed topical screening, whereas copies were also obtained for records later screened out on relevance, and those copies remain available for the bibliometric analysis reported here. Nineteen records were added during retrieval rather than by the original search, having surfaced through citation chasing while the located papers were being read, and they are marked in the screening matrix by discovery route so the distinction between systematic search and opportunistic addition stays visible. They are a disproportionately valuable nineteen, because they include Sherman's measurement-techniques report,

the Sohn, Cai and Vukovic source-localisation studies, the Patel and Vega Pasos blower-door comparisons, and a krypton-85 tracer study in a naturally ventilated agricultural building that exercises the cross-domain exception of Section 3.4.3 exactly as that clause anticipates.

The 49 records without an abstract are concentrated among conference proceedings and standards-body publications, particularly the ASTM special technical publications, where the indexed metadata is thin. This matters for the automated screening stage, because those records were classified on title and venue alone; they are flagged in the screening matrix for priority manual assessment and are the main source of classification uncertainty in the counts reported below.

4.1.1 Temporal Distribution

Figure 4.1 shows the distribution of publication years. The corpus spans 1967 to 2026. Its most important feature is that it is bimodal rather than a simple growth curve.

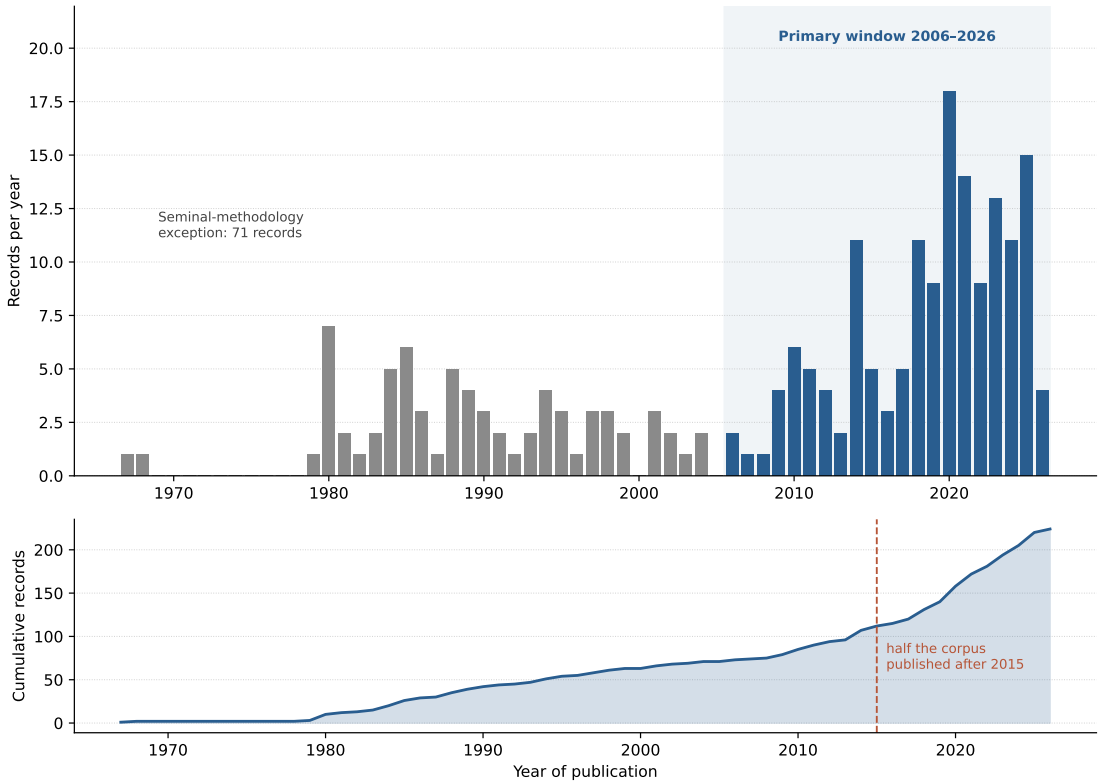


Figure 4.1: Publication activity across the retrieved corpus. The shaded region is the primary date window of Section 3.4, grey bars are records retained under the seminal-methodology exception of Section 3.4.3.

A first concentration runs from roughly 1980 to 1995 and contains the classical tracer-gas methodology, comprising the derivations of the decay, constant-injection and constant-concentration methods, the perfluorocarbon passive-sampling programme, and the measurement standards built on them. Seventy-one records fall before the primary window and are retained under the seminal-methodology exception, which is a substantial fraction and confirms that the exception clause was necessary rather than decorative, since applying the date window without it would have removed the papers that define the methods the thesis is about.

A second and larger concentration begins after 2015 and accelerates sharply from 2020; half the corpus was published after 2015. This is the low-cost sensing literature, and it is a different body of work with different authors, different venues and different concerns.

Between them lies a visible trough. The years 1996 to 2008 contribute only 21 records, the thinnest stretch of the entire span after the 1970s. The two concentrations are therefore not two phases of one continuous research programme but two largely separate literatures, roughly a generation apart. This observation recurs throughout the thesis, and it is the reason Chapter 5 and the low-cost sensing material are treated as distinct traditions that have only recently begun to converge.

4.1.2 Publication Venues

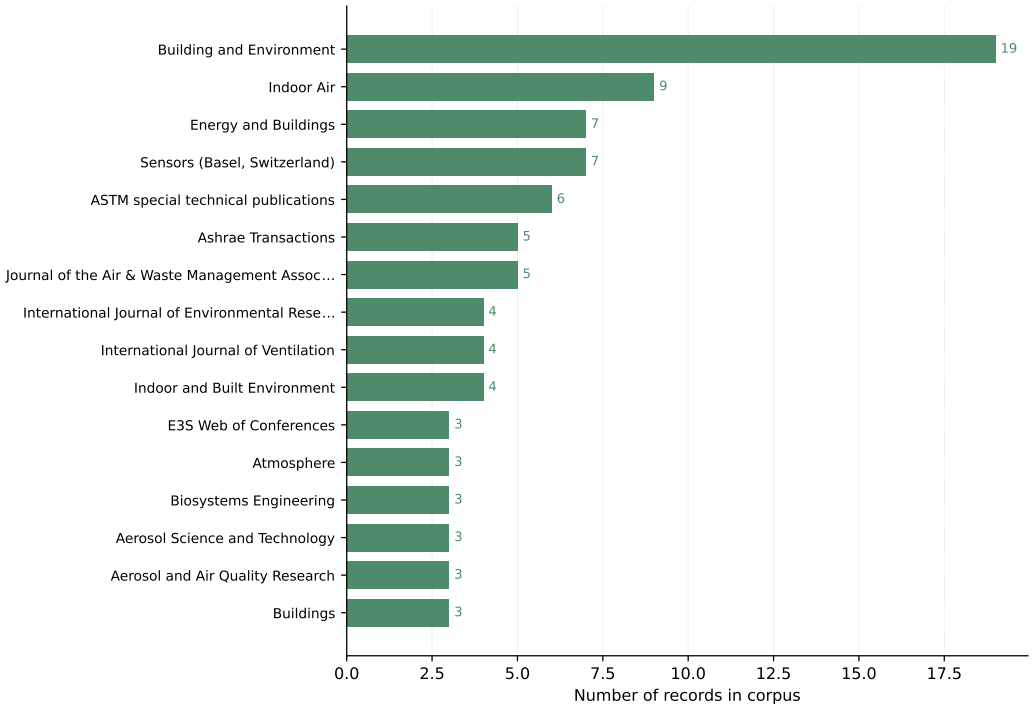


Figure 4.2: Principal publication venues in the corpus. Of the 226 records analysed, 178 carry an indexed venue, and those are spread across 91 distinct outlets of which 62 contribute a single paper. The chart shows only the leading venues, so the bars account for a minority of the corpus, and that dispersion is the result discussed in Section 4.1.

The corpus is markedly dispersed. The 178 records with an indexed venue are spread over 91 distinct outlets, and 62 of those outlets contribute exactly one paper. The largest single contributor, *Building and Environment*, accounts for 19 records, which is under nine per cent of the corpus. *Indoor Air* follows with nine, *Energy and Buildings* and *Sensors* with seven each, the ASTM special technical publications with six, and the *Journal of the Air & Waste Management Association* with five.

The composition of that leading group is itself informative. It mixes building-physics journals, an indoor-environment journal, an instrumentation journal and a standards series, which means the field has no disciplinary home. Work on the same physical problem is published to building engineers, to exposure scientists, to sensor developers and to standards committees, and these readerships cite each other only partially. This fragmentation is a practical argument for a review of this

kind, since a researcher following any single venue would see a distorted picture of the field, and the long tail of single-paper outlets means that citation chasing was necessary to reach material that keyword search alone did not surface.

4.1.3 Citation Impact

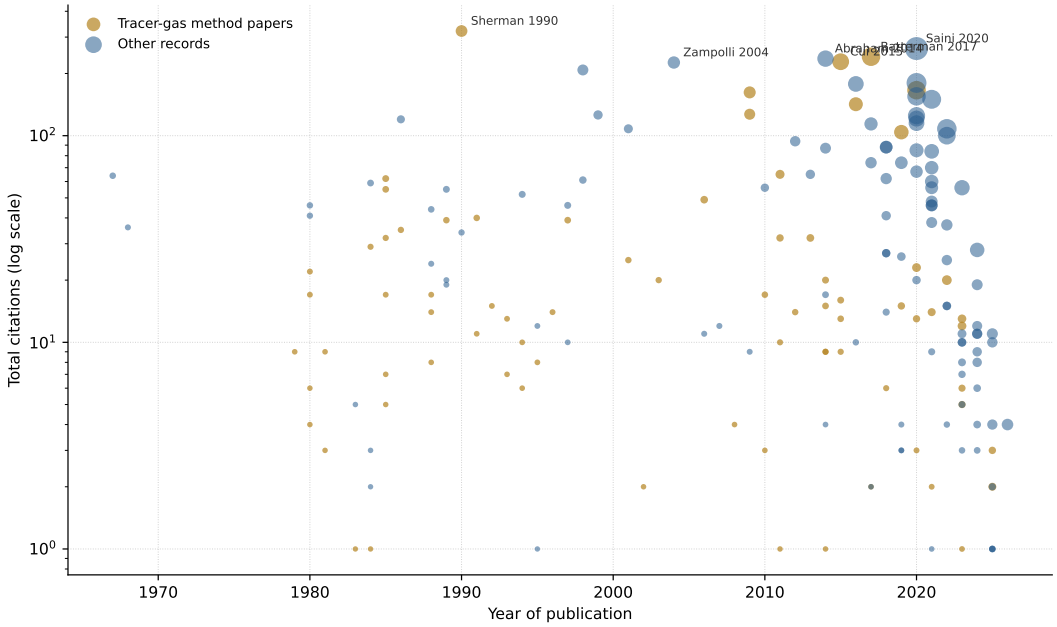


Figure 4.3: Total citations against year of publication, on a logarithmic scale. Marker area is proportional to citations per year.

Citation counts range from zero to 321, with a median of 11, and 45 records are uncited, most of them either recent or added by citation chasing and therefore carrying no citation metadata in the export. The most-cited items divide into two groups. The first comprises methodological reviews that consolidated the tracer-gas field, headed by Sherman's treatment of single-zone tracer techniques at 321 citations and the CO_2-based ventilation-rate methods review at 241. The second comprises recent surveys of low-cost IAQ monitoring, headed by a comprehensive review of monitoring systems at 264 citations and a cost-effective wireless sensor network study at 236.

Citation counts are used here only to characterise the corpus, and deliberately carry little weight in the selection score of Section 3.6. Two reasons argue against relying on them. Papers retained under the seminal exception have had decades to accumulate citations, so a threshold would systematically privilege age over quality. More importantly, the low-cost tracer-sensing work that is most directly relevant

to this thesis is recent and sparsely cited, and a citation-weighted selection would have discarded precisely the evidence the review exists to examine.

4.2 Distribution Across the Keyword Taxonomy

Classifying every record against the taxonomy of Section 3.9 gives the branch totals in Table 4.1. Three records match no branch at all, which is an acceptable coverage rate and indicates the vocabulary is not missing a substantial theme.

Table 4.1: Corpus coverage by taxonomy branch

Branch	**Records**	**%**	**Most frequent leaf**
Tracer gases (targets)	172	76	Tracer gas methods (98)
Application domain	134	59	Indoor air quality (97)
System characteristics	134	59	Low-cost sensing (74)
Ventilation parameters	111	49	Ventilation rate (37)
Specific applications	60	27	Airflow pathway analysis (18)
Sensor technologies	34	15	Gas chromatography (9)

The distribution is uneven in a way that is diagnostic rather than incidental.

The tracer-gas branch dominates, as the search strategy intends. Within it, carbon dioxide appears in 75 records and sulfur hexafluoride in 31, while the alternatives the supervisor asked to have investigated are thinly represented, with hydrocarbon tracers appearing in five records, refrigerant tracers in three, and noble-gas tracers in one. The literature on alternatives to SF_6 and CO_2 is therefore genuinely small, not merely hard to find, and this scarcity is a finding in its own right rather than a limitation of the search.

The sensor-technologies branch is the thinnest at 15 per cent, and this is the most revealing number in the table. Gas chromatography leads with nine records, NDIR follows with eight, and photoacoustic detection, photoionisation and electrochemical sensing appear in three, two and two records respectively. Most of the corpus discusses measuring a tracer without specifying the detection principle used to measure it. For a thesis about the development of low-cost sensors this is a structural gap in the evidence base, since the literature is rich on ventilation physics and increasingly rich on deployment architectures, but comparatively silent on the transduction step that determines whether a low-cost instrument can substitute for a reference one.

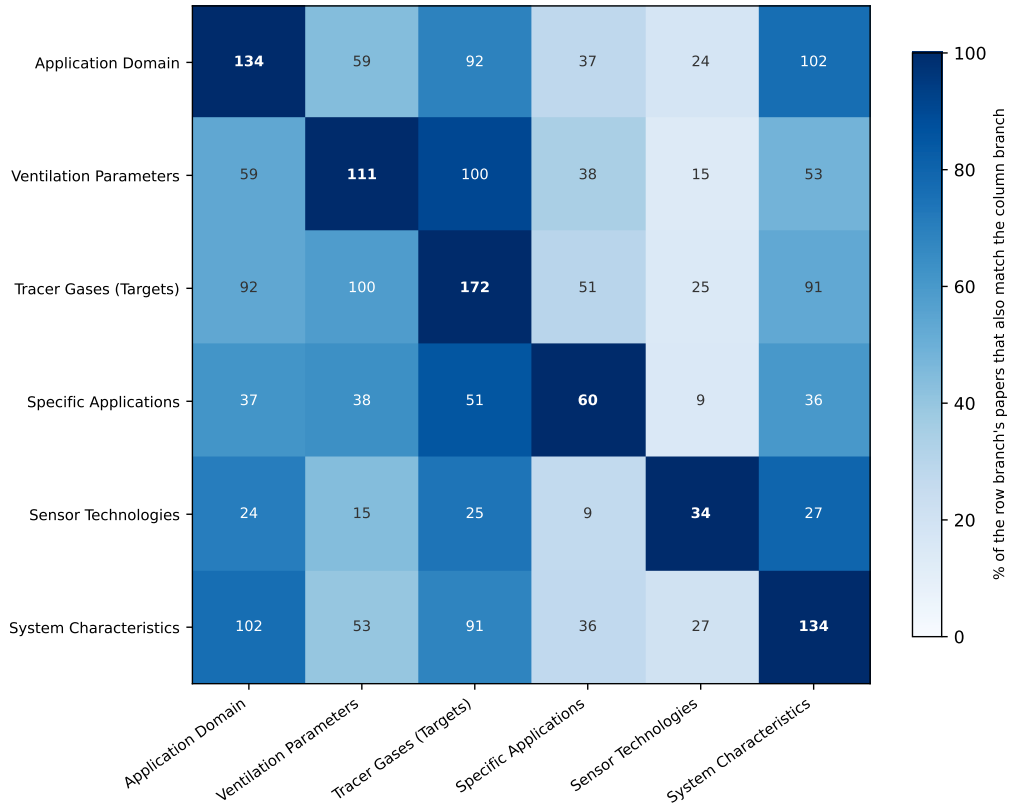


Figure 4.4: Co-occurrence between taxonomy branches. Cell values are record counts and shading gives the percentage of the row branch's records that also match the column branch.

The co-occurrence matrix in Figure 4.4 shows how the branches combine. Tracer gases and ventilation parameters co-occur in 100 records, which is expected given that tracers are the instrument by which ventilation parameters are determined. Application domain and system characteristics co-occur in 102 records, reflecting the deployment-oriented low-cost literature. The weak cells are the informative ones, because sensor technologies co-occur with specific applications in only nine records, meaning that studies naming a detection principle rarely also address a concrete diagnostic problem such as envelope leakage or source localisation.

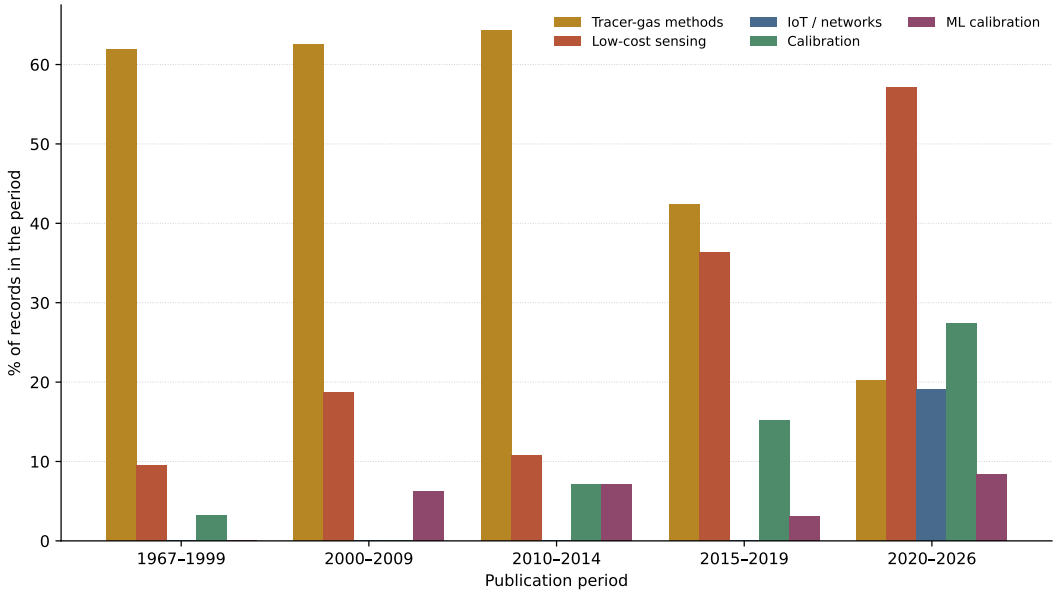


Figure 4.5: Share of records in each period matching selected themes, showing the displacement of tracer-gas methodology by low-cost sensing concerns.

Figure 4.5 tracks how emphasis has moved. Tracer-gas methodology dominates the pre-2000 periods and declines steadily as a share of output thereafter, while low-cost sensing, networked deployment and machine-learning calibration are almost absent before 2010 and rise sharply after 2015. The substitution is close to complete, in that the modern literature has largely stopped asking how accurately a ventilation rate can be measured and started asking how cheaply and how continuously a concentration can be logged. These are not the same question, and the gap between them is where this thesis is situated.

4.3 The Evidence Funnel

The preceding observations converge on a single quantity, which is how much of the retrieved literature actually addresses the thesis topic, that being neither tracer-gas methodology alone nor low-cost sensing alone but their intersection. Figure 4.6 reports it.

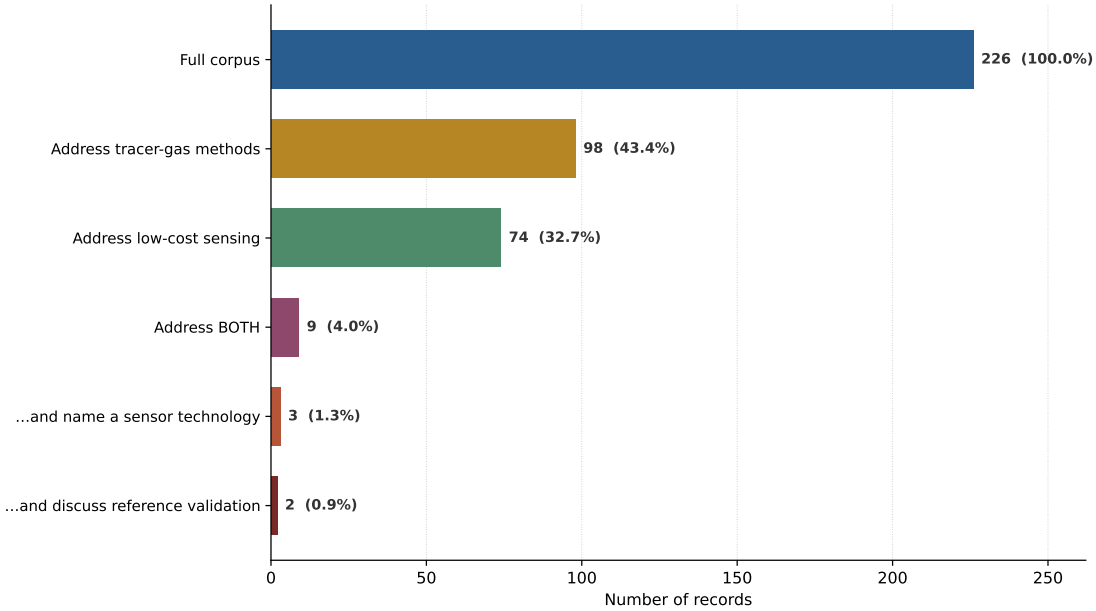


Figure 4.6: Successive restriction of the 226-record analysis corpus to the intersection of tracer-gas methodology and low-cost sensing. The denominator is the analysis corpus rather than the 173-record synthesis corpus, because the funnel measures how much of the retrieved literature works at the intersection, and removing records excluded for reasons unrelated to topic, such as an unobtainable full text, would understate the field rather than describe it.

Within the 226-record analysis corpus, 98 records address tracer-gas methods and 74 address low-cost sensing, but only nine address both. Of those nine, three name the sensor technology used, and two discuss validation against a reference instrument. Nine records out of 226 is 4.0 per cent of the corpus, and two records is under one per cent.

That this ratio did not move when the corpus grew is itself worth noting. Adding nineteen papers found by following citations out of the retrieved literature, which is the search route most likely to surface work at an intersection because it follows the reasoning of authors already working there, produced seven more tracer-gas methodology papers and no additional papers at the intersection. The gap is therefore a property of the field rather than an artefact of the search strings.

This is the central result of the search, and it warrants stating plainly. The literature on measuring ventilation with tracer gases is mature and the literature on low-cost indoor air quality sensing is large and growing quickly, but the body of work that combines them, asking whether an inexpensive sensor can perform a tracer-gas measurement to a standard that would let it replace a reference instrument, consists of a handful of studies. The publication years of those nine records reinforce the

point made in Section 4.1, since three fall in the classical period (1979, 1981, 1988) and are comparisons of detector cost rather than low-cost sensing in the modern sense, while the genuinely relevant modern work is dated 2006, 2010, 2017, 2023 and two in 2025. The convergence is recent and it is still in progress.

Three consequences follow, and they shape the remainder of the thesis. First, the review cannot be a meta-analysis, since with nine studies at the intersection there is nothing to pool statistically, and the appropriate treatment is close critical reading of each. Second, the thesis must draw its physical criteria from the classical literature and its engineering constraints from the low-cost literature, because neither supplies both, which is why the tracer-selection criteria established in Section 2.4 are applied as an external framework rather than adopted from the low-cost sensing work. Third, the scarcity of validation evidence, amounting to two records, means that claims about the field readiness of low-cost tracer sensing should be treated as provisional throughout, and this caution is carried into the assessment of individual platforms.

4.4 Limitations of the Corpus Analysis

Four qualifications apply to the figures in this chapter.

The classification is based on title and abstract text, so it detects the vocabulary a paper uses rather than its full content. A study that applies NDIR detection without naming the principle in its abstract is not counted under that leaf. The sensor-technologies total of 34 records is therefore a lower bound, and the true figure will rise as full-text assessment proceeds. The direction of this bias is known, which is why the branch is described above as thinly *reported* rather than simply absent.

Forty-nine records carry no abstract and were classified on title alone, which compounds the same effect. These are disproportionately the older standards and proceedings material, so the classical portion of the corpus is more likely to be under-classified than the modern portion.

Seventeen records were excluded from the scored matrix for want of a full text, comprising fifteen for which no copy could be obtained at all, and two whose downloaded file contained no assessable article, one being a single-page conference poster and the other an abstract-only conference submission. Two of the fifteen are recorded as permanently unavailable, having proved untraceable through the library, the publisher and the secondhand market alike. Because methodological soundness and evidence strength cannot be judged without reading the work, these records are not scored, not cited and not carried into the bibliography. This is a retrieval limitation rather than an analytical one, since the eight papers in the corpus that arrived as scanned page images did hold complete articles and were read and scored

in the ordinary way, whereas a document that was never obtained cannot be read by any means. The unscored records are concentrated in the pre-2000 conference literature, with twelve of the seventeen carrying no DOI, which is consistent with the venue fragmentation described in Section 4.1. Their absence biases the corpus analysis in the same direction as the missing abstracts, by under-representing the classical material relative to the modern.

The counts are not mutually exclusive. A record matching several leaves is counted under each, so branch totals sum to more than 226 and percentages do not add to 100. The co-occurrence matrix in Figure 4.4 is provided precisely so that the overlap is visible rather than hidden.

Finally, the corpus reflects the search strategy of Chapter 3 and inherits its limitations, in particular English-language indexing and the under-representation of grey literature. Manufacturer validation reports and unpublished field trials are likely to contain relevant low-cost tracer-sensing evidence and are largely invisible to this search. Given that the intersection identified in Section 4.3 contains only nine records, the addition of even a small amount of grey literature could change the picture materially, and this is the single most consequential limitation of the review.

Chapter 5

Tracer Gases for Ventilation Measurement

Chapter 2 established the mass balance from which every tracer measurement follows, and set out the four criteria a tracer must satisfy for that measurement to be valid. This chapter applies those criteria to the candidate gases reported in the corpus.

The gases are treated in three families rather than individually, because the constraints that determine suitability act at the family level. The halogenated reference tracers of Section 5.2 share extreme detectability and a high environmental burden. The occupant and combustion gases of Section 5.3 share a variable indoor background and, in two cases, a physiological source that removes the need for a cylinder. The lower-burden alternatives of Section 5.4 are grouped by what motivates them, since each was proposed to relieve the cost or environmental penalty of the first family and each pays for that relief somewhere else. Section 5.5 then compares all of them against the four criteria directly, Section 5.6 treats the error sources that are independent of gas identity, and Section 5.7 states the selection guidance the evidence supports.

Figure 5.1 summarises the position the chapter argues for. The two axes are the properties that dominate selection in practice, namely how readily a tracer can be distinguished from indoor background and what it costs to use in environmental, safety and logistical terms. No candidate occupies the favourable corner, and that is the reason tracer selection remains a compromise rather than a solved problem.



Figure 5.1: Tracer gases positioned by detectability against indoor background on the horizontal axis and by burden of use, combining environmental, safety and logistical cost, on the vertical axis. Positions are qualitative and follow the assessments made in this chapter. The reference-grade tracers achieve detectability at high burden, carbon dioxide achieves low burden at poor detectability, and the alternatives occupy the middle without resolving the trade-off, so that no candidate occupies the favourable lower-right corner. That absence is the reason tracer selection remains a compromise rather than a solved problem.

5.1 The Selection Problem

The choice of tracer is frequently reported as a preference rather than as a decision with stated grounds. Treating it as a decision requires being explicit about what is being traded.

Detectability against background determines how much tracer must be released. A gas absent from indoor air can be measured at very low concentration, which permits a small release, a short mixing period and a rapid test. A gas already present at variable concentration must either be released in quantity or have its background measured and subtracted, and the subtraction carries its own uncertainty into the result through the $C - C_{out}$ term of Equation 2.3.

Burden of use combines three constraints that are usually reported separately but bind together in practice. Environmental impact now restricts the quantity of

some gases that may be released at all, safety determines whether a test may run in an occupied building, and logistics covers cylinders, permits, analytical equipment and the expertise needed to operate it.

The two are correlated in the wrong direction. The gases that are easiest to detect, the halogenated compounds, are those with the highest environmental burden, because the same chemical stability that keeps them conserved during a measurement keeps them in the atmosphere afterwards. This is not an accident of the current inventory but a consequence of the property that makes a good tracer, and it explains why the search for alternatives described in Section 5.4 has proved harder than it appears.

5.2 Halogenated Reference Tracers

Sulfur hexafluoride and the perfluorocarbon family are the reference standard against which other tracers are judged. They share the properties that define the category, which are negligible indoor background, high chemical stability, detectability at extremely low concentration by electron-capture methods, and a substantial global warming potential.

5.2.1 Sulfur hexafluoride

SF_6 has been used for air-change measurement, duct and air-handler testing, re-entrainment studies, laboratory-hood containment, mine and industrial ventilation, and mixing experiments [4], [9], [11], [25]. Its standing derives from a very low atmospheric background, strong analytical detectability and compatibility with the decay, constant-injection and pulse designs alike. Early work described it as close to an ideal tracer on the grounds that very small concentrations can be measured and it is not a normal constituent of indoor air [26].

The principal analytical routes are gas chromatography with electron-capture detection, photoacoustic infrared spectroscopy, infrared absorption instruments and portable direct-reading analysers. Electron-capture detection reaches the lowest concentrations; photoacoustic instruments provide real-time measurement; portable infrared analysis made SF_6 usable in field studies where laboratory chromatography was impractical [26], [27], [28]. For duct measurements, injection and sampling must be arranged so that the gas is adequately mixed across the section before the measurement plane [11].

Three limitations qualify this position. The first is environmental, because SF_6 is a potent greenhouse gas, and this has driven sustained effort to reduce or replace its use in standardised tests [29], [30], [31]. The second is that it is not perfectly conserved, as porous construction materials can take up and later release the tracer, so the assumption that SF_6 remains within the measured volume does not hold

universally. The third is physical, in that SF_6 is a heavy gas, and where release momentum or buoyancy contributes to its transport the measurement describes the tracer's behaviour rather than the airflow, in the manner quantified by Gao and colleagues [1].

Measurements are additionally vulnerable to incomplete mixing, sampling-location bias, and the disturbance created by the mixing fans used to address the first of these. In calm conditions such fans can alter the infiltration being measured [32]. A high-quality detector does not by itself produce a high-quality ventilation estimate.

SF_6 therefore remains the strongest choice where the priority is sensitivity, comparability with historical data, or a short and carefully controlled test, and a poor choice where the measurement is long, occupied, or environmentally constrained.

5.2.2 Perfluorocarbon tracers

Perfluorocarbon tracers address the case SF_6 handles worst, which is the multiday, multizone field measurement. Rather than a single gas, the method uses a family of chemically distinct fluorinated compounds, among them perfluoromethylcyclohexane, perfluorodimethylcyclohexane, hexafluorobenzene, octafluorotoluene, perfluoroallylbenzene and octafluorobenzene [33], [34], [35], [16], [36]. Deploying different compounds in different zones is what makes multizone measurement possible, because each zone is tagged separately and interzonal as well as outdoor exchange can be resolved from a single deployment.

The Brookhaven Air Infiltration Measurement System established the practical form. Miniature passive sources emit tracer vapour at a known rate while passive adsorption tubes integrate concentration over hours or days; the tubes are then thermally desorbed and analysed by chromatography. The method reaches part-per-quadrillion concentrations and operates over periods from roughly twelve hours upward, with longer deployments usual in field surveys [34], [35], [10]. Later variants used passive dosers with carbon molecular-sieve cartridges followed by solvent extraction and mass-spectrometric analysis; one modified method reported 24-hour detection limits of approximately 0.17–0.26 $\mu g\,m^{-3}$ at passive sampling rates of 10–14 $mL\,min^{-1}$ [16].

The limitations are practical rather than analytical. The method integrates, so it cannot resolve airflow that changes quickly. Analysis requires chromatography and careful sample handling, which places the cost in the laboratory rather than in the field hardware, a distinction that matters when such methods are described as low-cost. Source emission rates are temperature sensitive, and passive sampling adds uncertainty through recovery efficiency, storage, contamination risk and calibration drift [37], [38]. Perfluorocarbons are also greenhouse gases, so the environmental objection that applies to SF_6 applies here in a milder form.

Leaderer and colleagues characterised these error terms individually in a controlled chamber, and the resulting picture is unusually complete for a passive method [39]. The sampling hardware performed better than expected, since the relative standard deviation between paired passive samplers was 1.9 ± 1.0 per cent, which the authors took as sufficient to make duplicate sampling unnecessary in the field, and sampler orientation contributed at most a 2 to 3 per cent bias at the low air velocities typical of dwellings. Temperature sensitivity of the source proved manageable, with a time-weighted average temperature accounting for fluctuations as large as 8 °C.

The dominant error is structural rather than instrumental, and it survives any improvement to the sampler. A ventilation rate is proportional to the average of the *reciprocal* concentration, whereas a passive integrating sampler measures the reciprocal of the *average* concentration. The two are equal only if the concentration is constant, so whenever ventilation varies during the integration period the method returns a biased result, estimated at 3 to 6 per cent under typical conditions and reaching 10 per cent when the ventilation rate varied threefold during the measurement. Leaderer and colleagues judged this tolerable for the convenience the technique offers, which is a defensible position, but it is a bias rather than a random error and it does not average out over repeated deployments. It is also a specific instance of the general point developed in Chapter 9, which is that the limitation lies in the inference from concentration to flow rather than in the detection of the concentration.

5.2.3 What the family shares

Both tracers achieve their detectability through chemical inertness, and both incur their environmental burden through the same property. Neither can be made low-burden by using less of it, because the detection advantage is what permits small releases in the first place and the release quantity is already near the practical minimum. Displacing this family therefore requires a different detection principle rather than a different quantity, which is the observation that motivates Section 5.4 and Chapter 7.

5.3 Occupant and Combustion Gases

The second family is defined by a shared difficulty, in that each gas has a substantial and variable indoor background, so the C_{out} term is neither small nor stable. Two of the three compensate by removing the need for a release altogether.

5.3.1 Carbon dioxide

Carbon dioxide is the most practical tracer for occupied buildings, and the reason is logistical rather than analytical. Occupants generate it, so no cylinder, permit or release equipment is required; it is measurable with inexpensive non-dispersive infrared sensors; and it presents no exposure hazard at the concentrations involved [2], [13], [40].

It is used in three distinct ways, which are frequently conflated and have different error structures.

1. **Occupant-generated**, where the occupants are the source and the generation rate must be estimated;
2. **Injected**, where a known quantity is released for a decay or constant-injection test, giving a controlled source at the cost of the logistical advantage;
3. **Ambient difference**, where indoor and outdoor concentrations are compared to estimate outdoor-air delivery or recirculation.

Only the second gives a controlled source term. The first, which is the common case in the low-cost literature, replaces a known G with an estimated one, and the estimate depends on body size, activity, age and occupancy pattern [14], [41]. Where a direct comparison has been made the results are encouraging, since in a residence study carbon dioxide and SF_6 produced statistically similar air-change results, with carbon dioxide simpler and less expensive [42].

The three uses have been compared quantitatively, and the ranking is consistent enough to guide selection. Liang and colleagues measured air change rates in 85 naturally ventilated dormitory rooms under open- and closed-window conditions and propagated the uncertainty of each method through to the result [43]. The mean relative standard uncertainties were 68 per cent for the build-up method, 38 per cent for the steady-state method and 29 per cent for decay. More useful than the ordering is the reason for it. For build-up and steady state, the largest single contributor was the human CO_2 generation rate, which is the estimated G of the first case, and that confirms that the dominant error in occupant-based measurement is the source term rather than the sensor. For the decay method, where the occupants have left and G is zero, the largest contributor was instead the outdoor CO_2 concentration. That is the background term C_{out}, and its dominance is exactly what Equation 9.2 predicts once the source term is removed, because with G eliminated the remaining error migrates to the background estimate and to the tail of the decay.

The practical consequence is that switching from build-up to decay does not remove the uncertainty so much as relocate it, from a quantity that must be estimated from the occupants to one that must be measured outdoors. A low-cost

deployment that omits an outdoor node has therefore left its dominant error source unmeasured.

Two limitations dominate. The background is substantial and variable, so background correction is essential rather than optional, and an error in it propagates directly through Equation 2.3. The measurement is also unusually sensitive to mixing, since under displacement ventilation, natural ventilation or stratified conditions a wall-mounted sensor can differ substantially from both the breathing-zone and the room-average concentration [44], [45], [46]. A carbon dioxide threshold can therefore be useful for screening and control without constituting an air-change measurement, a distinction pursued in Section 6.2.

5.3.2 Nitrous oxide

Nitrous oxide has a long history in industrial hygiene, building ventilation and containment testing, measured by infrared spectroscopy against a low natural indoor background [26], [47]. Kalliokoski and colleagues judged continuous nitrous oxide monitoring the most practicable option for ventilation testing under their field conditions [26].

Its position in the literature is genuinely divided, and the division is instructive. In controlled mixing and uniformity tests it has reproduced SF_6 behaviour closely while reducing global-warming impact [30], whereas in laboratory-hood containment testing it did not correlate strongly enough with SF_6 across different hood geometries and face velocities to serve as a drop-in replacement [29]. The conclusion that follows is methodological rather than chemical, in that substitution must be validated for the specific geometry and test standard rather than inferred from molecular similarity.

Practical disadvantages include infrared interference from water vapour and carbon dioxide, reduced accuracy at low concentration, occupational exposure limits that constrain use in occupied buildings, and solubility and surface absorption effects that become significant at low concentration in airtight spaces [8]. Its environmental impact is lower than that of SF_6 but not negligible.

In large single-cell industrial spaces the practical choice has tended to fall on SF_6 regardless, because electron-capture detection provides the sensitivity a large volume demands. Simons and Waters built an automated six-channel decay system for factory buildings around SF_6 with a pulse-modulated electron-capture chromatograph on each channel, tuning the detector to SF_6 specifically in order to suppress interference from other electrophilic compounds present in industrial atmospheres [48].

5.3.3 Carbon monoxide

Carbon monoxide appears in early building and industrial-ventilation studies, usually measured by non-dispersive infrared instruments, and has been compared directly against SF_6 in chambers and test houses [27]. Its advantages are instrumentation familiarity and, relative to SF_6, a low environmental burden.

Its disqualifying property is toxicity. A gas that is itself a serious indoor pollutant is unsuited to occupied-building work unless concentrations are tightly controlled and exposure prevented, which removes the practical advantage that motivated its use. It is best regarded as a historical option retained in the literature for comparison rather than as a current candidate.

5.4 Lower-Burden Alternatives

The third family is defined by intent rather than chemistry. Each candidate was proposed to relieve the cost or environmental penalty of the halogenated tracers, and examining them together makes visible what each concedes in exchange.

5.4.1 Isobutene and photoionisation detection

Isobutene is the best-developed of the lower-cost alternatives and the only one in this family with an occupied-building reference comparison. It is measured with a photoionisation detector, substantially cheaper than the photoacoustic system used for SF_6 in the same study. In new apartment buildings, isobutene and SF_6 produced closely agreeing ventilation rates near $1\,h^{-1}$ [49].

The qualifications are specific. Because a photoionisation detector responds to volatile organic compounds generally, the tracer concentration must stand clear of the indoor background; the study estimated that a level at least ten times background was required to hold the background contribution near ten per cent [49]. Isobutene reacts with ozone, although this was judged negligible at the indoor concentrations encountered. Flammability requires consideration even at working concentrations far below the explosive range.

Isobutene therefore suits short, lower-cost decay tests where the volatile organic background has been characterised and the space is unoccupied or controlled. It is not chemically unambiguous in the way SF_6 is, and the trade it makes is selectivity for cost.

5.4.2 Refrigerant tracers

R134a has been used where the objective is dense spatial mapping rather than a single accurate air-change figure. A wireless grid detected it with tin-oxide

semiconductor sensors over a nominal 5–100 ppm range, achieving a mean air-change error of approximately eight per cent against a photoacoustic reference and reducing the sampling interval from minutes to about ten seconds [50].

The contribution is the sampling density rather than the tracer. Inexpensive nodes permit simultaneous measurement at many locations, which no affordable reference instrument allows. The concessions are correspondingly specific, in that tin-oxide sensors are cross-sensitive, passive diffusion is slow in stagnant air, agreement degrades outside the useful concentration range, and wireless data loss can compromise a test. R134a is itself a greenhouse gas, so small release quantities reduce but do not remove the environmental objection [50]. The platform is examined in full in Section 7.4.

One further caveat applies to any refrigerant used as a tracer, and it is the criterion of low background rather than any property of the detector. Fisk and colleagues, screening candidate halocarbon tracers across several buildings, found one in which the background concentration of R-12 was approximately 225 ppb, high enough to disqualify that compound as a tracer in that building; the source was a large refrigerant leak in the mechanical system [51]. The buildings in which a refrigerant background is most likely to be elevated are precisely those with ageing or faulty plant, which are also the buildings most likely to warrant a ventilation investigation. A background check before the release is therefore not optional, and it is one measurement a low-cost node can perform for itself.

5.4.3 Light hydrocarbons

Methane has been compared directly against the reference tracers in an occupied-scale intercomparison. Grimsrud and colleagues ran concurrent decays in a single-storey research house and found air-change rates measured with SF_6 slightly larger than those measured with the lighter gases, the ratio being 1.16 ± 0.09 against methane and 1.09 ± 0.09 against nitrous oxide, with an overall mean of 1.10 ± 0.10 [52]. The authors could not explain the offset by molecular diffusion or by settling of the heavier tracer, and concluded that although the difference appears real it sits inside the 5 to 10 per cent uncertainty of any single tracer measurement and should not by itself decide between the gases. Methane is therefore a defensible substitute on agreement grounds, and the objection to it is flammability rather than accuracy. Ethylene was tested as an SF_6 replacement for laboratory-hood performance testing, but its relationship with SF_6 proved insufficiently consistent to recommend [53]. Propane appears in recent low-cost infrared sensor development as a gas the platform might be adapted to, though the building-ventilation evidence is far thinner than for SF_6 [54].

The case for light hydrocarbons is clearest where occupancy is the obstacle.

Rooney and colleagues, measuring room-level ventilation across schools and universities, used methane and ethane alongside CO_2 for two stated reasons that follow directly from the selection criteria of Section 2.4, namely that neither gas is emitted by humans, so occupants elsewhere in the building cannot perturb the test, and that both have a low and stable natural background [55]. Methane was released at one per cent in nitrogen to exclude any ignition hazard, raising the room concentration to 10–20 ppm before the decay was followed. The comparison with CO_2 in the same rooms is informative in both directions, since measured backgrounds agreed with outdoor readings to within the sensor's resolution of 50 ppm in all but one classroom, so for these rooms the CO_2 background problem was small in practice, although establishing that required the hydrocarbon measurement to check it against.

Two further observations from that work transfer to low-cost practice. Plotting the decay in log-linear space made transient changes in ventilation visible as breaks in slope, which is a diagnostic available to any platform that records the full decay rather than fitting it automatically. And in a quarter of the classrooms the controlled-release result disagreed with a balometer measurement of supply flow by as much as $6.2\,\mathrm{h}^{-1}$, which is a reminder that a disagreement between two methods locates a problem without identifying which of them holds it.

The common constraint is flammability, which restricts use in occupied buildings and imposes handling requirements that erode the cost advantage motivating the substitution.

5.4.4 Noble-gas tracers

Radioactive noble gases including argon-41 and krypton-85 appear in the historical literature where radioactive detection or difficult large-space problems were involved [7]. Krypton-85 has been used for air-exchange measurement in naturally ventilated agricultural buildings [56], [57]. These methods are sensitive and effective in large or awkward spaces, but source management, regulation, specialised equipment and operational burden rule them out for routine occupied-building work.

Stable krypton is the more relevant modern case. Krypton-84 was validated as a possible replacement for SF_6 in standardised containment and protection-system testing, with results reported as equivalent under the conditions tested [31]. It avoids the greenhouse-gas profile of SF_6 entirely, which is the strongest environmental case in this family, but requires specialised detection and standard-specific validation before it can be adopted.

Helium, hydrogen, oxygen and acetone appear in historical reviews of gases used or considered for airflow studies [7], and should not be read as equally established. Helium is inert but differs greatly from air in density, with cost and supply constraints. Hydrogen detects well in some systems but is flammable.

Oxygen is abundant in background air and cannot be distinguished without isotopic or instrumental separation. Acetone is detectable with volatile organic instruments but is reactive and common enough indoors to require careful background control.

5.4.5 Passive volatile organic tracers

The most recent direction seeks a tracer that can be released passively and captured with the same samplers already used for indoor volatile organic measurement, aligning the ventilation and pollutant measurements in time and removing the need for a separate active analyser [18]. If successful this changes the cost structure rather than the tracer, since the marginal cost of the air-change measurement falls to near zero when it rides on an analysis the study is performing regardless.

Two candidates illustrate the difficulty. 2-butoxyethyl acetate was investigated because it had low observed background, could be captured by commercial passive samplers, and was relatively low in chronic toxicity, but passive-source testing produced unstable emission, and the compound was subsequently discarded when hygroscopic behaviour proved to make the source rate too sensitive to moisture [17]. Deuterated decane is the more promising successor, since isotope labelling provides analytical distinguishability against a low natural background while leaving transport behaviour essentially unchanged, which is precisely the property a conservative tracer requires [18], [58]. Source design and cost have so far prevented routine use.

The rejected candidate is as informative as the successful one. A tracer must be evaluated as a complete measurement chain, covering source, room transport, passive sampling, extraction, analysis and uncertainty, and a compound that is analytically measurable can still fail because its release rate is not stable. This is the same conclusion the sensor literature reaches from the opposite direction in Section 6.7.

5.5 Comparative Assessment

The preceding sections treat each family on its own terms. This section compares all candidates directly against the four criteria of Section 2.4, namely physical and transport properties, background and detectability, health, safety and environmental burden, and practical use in buildings, so that the trade-offs are visible in one place. Table 5.1 presents that comparison.

Table 5.1: Analytical Comparison of Major Tracer Gases Against Key Evaluation Criteria

Gas / Molecule	Physical & Transport	Background & Detectability	Health, Safety, Environment	Practical Use in Buildings
SF_6	Prone to material interactions and mixing bias [29].	GC-ECD, PAS, IR. Very sensitive; low background [4], [26].	Very high greenhouse impact.	Short-term reference, decay, ducts, mixing.
PFTs	Subject to source-rate, temperature, zone, and steady-state errors [37], [38].	Passive adsorption plus GC or GC-MS.	—	Long-term passive and multizone field studies; low burden, scalable [33], [35].
CO_2	Susceptible to mixing and placement bias [2], [45].	NDIR. Variable background and indoor generation [2], [45].	Safe for standard occupancy.	Occupied rooms, schools, DCV, low-cost networks; cheap, available [40], [59].
N_2O	Transport is not universally equivalent to SF_6 [29], [30].	IR spectroscopy. Low background; suffers from water-vapour interference [26].	—	Workable for continuous infrared-monitored ventilation and selected substitutions.
Isobutene	Subject to ozone reaction.	PID. Can be masked by general VOC background.	Flammability risk; lower environmental load than SF_6 [49].	Lower-cost short-term decay tests.
R134a	Constrained by diffusion limits [50].	MOS or specialized gas sensors; subject to cross-sensitivity.	Known greenhouse impact.	Dense spatial mapping; wireless, fast, inexpensive spatial coverage [50].
CO	Exhibits detectable decay [27].	NDIR. Utilizes familiar instrumentation.	High toxicity makes occupied use unattractive [27].	Historical chamber and industrial studies.

Continued on next page...

Table 5.1: Analytical Comparison of Major Tracer Gases (Continued)

Gas / Molecule	Physical & Transport Properties	Background & Detectability	Health, Safety & Environment	Practical Use in Buildings
Methane or ethylene	Methane agrees with SF_6 to within about 16% in concurrent decays [52], while ethylene proved insufficiently consistent for hood testing [53].	Hydrocarbon detector or direct-reading instrument.	Presents flammability risks [53].	Specialized comparative or hood studies; potentially inexpensive.
Krypton	Highly stable noble gas.	Specialized noble-gas detection required.	Avoids greenhouse-gas use [31].	Specialized or standardized replacement tests; specialized handling, limited IAQ evidence.
Deuterated compounds	Presents source-design challenges.	GC-MS or isotope-sensitive analysis. Low natural background and analytical distinction [18], [17].	—	Emerging passive VOC-compatible tests; high cost.

5.5.1 The comparison performed directly

The matrix above is assembled from studies that each examined one tracer. One study in the corpus applied several methods to the same building over several months and so tests the comparison rather than compiling it. Dols and Persily, working in a large mechanically ventilated office building, ran automated SF_6 tracer decay alongside CO_2 measurement and direct airflow measurement at the outdoor-air intakes [60].

The reported uncertainties separate the candidates cleanly. SF_6 concentrations determined by gas chromatography with electron capture detection carried about 10 per cent uncertainty, and the air change rates derived from them about 10 per cent under the assumption of perfect mixing. CO_2 measured by infrared absorption carried an uncertainty of 12.5 ppm, while the same quantity read from a colorimetric detector tube carried 33 ppm, set purely by the resolution of the graduated markings. Direct airflow measurement at the intakes was the most precise of the three at better than 3 per cent from instrument error alone, but that

figure excludes error from traverse positions not conforming to the recommended procedure, which is where the real uncertainty of a duct traverse usually lies.

Two general points follow. First, the ranking of methods by uncertainty depends on which error terms are counted, and the method with the smallest instrument error here is the one whose procedural error is least well characterised. Second, the difference between the infrared analyser and the detector tube, measuring the same gas in the same building, is a factor of nearly three, which shows that the choice of tracer does not determine the achievable uncertainty on its own; the detection method does at least as much. This is the comparative counterpart of the argument developed throughout Chapter 6.

5.6 Cross-cutting Sources of Uncertainty

5.6.1 Mixing and Sampling Representativeness

The most persistent error is treating a room as one well-mixed zone when it is not. Stratification, displacement ventilation, open doors, furniture, local sources, and interzonal flows can make a single concentration unrepresentative. PFT studies show that even whole homes can behave as several zones, while classroom studies show that CO_2 readings depend strongly on sensor height and location [38], [44], [45].

5.6.2 Time-varying Airflow

Natural ventilation is driven by wind and buoyancy, so air-change rate can vary during a test. Repeated decay tests in dwellings showed large variability, with coefficients of variation reported from 20% to 64% in one study [14]. A single tracer result should therefore be treated as a time-specific estimate unless the test duration and weather conditions support a longer-term interpretation.

5.6.3 Source Behaviour

Active injection is easier to define but more intrusive. Passive sources are less intrusive but require calibration of emission rate. Temperature, humidity, cap design, vial geometry, and material selection can alter the dose rate [16], [37], [17].

5.6.4 Chemistry and Building Materials

A tracer should not react with indoor oxidants, adsorb strongly to surfaces, permeate walls, or be removed by filtration during the measurement. That these interactions are reported even for SF_6, a gas commonly described as inert, shows that chemical

stability in the abstract is not the same as conservation inside a real building envelope.

5.6.5 Instrumentation and Data Reduction

Detector choice changes the usable concentration range, time resolution, cost, and selectivity. GC methods are sensitive but often laboratory-based; PAS and IR instruments provide real-time data but can be expensive; PID systems are fast but respond to multiple VOCs; MOS devices are inexpensive but require calibration and may be cross-sensitive [26], [49], [50].

5.7 Selection Guidance

No tracer is universally optimal, so selection must be matched to the measurement. The guidance below follows from the comparative matrix and states the conditions under which each candidate is defensible.

- Use SF_6 when the study needs a well-established short-term reference method and environmental release can be minimized and managed.
- Use PFTs when the goal is a multiday or multiroom field average with minimal disturbance.
- Use CO_2 when the study is occupied, low-cost, or focused on screening, demand-controlled ventilation, or dense monitoring.
- Use N_2O only after confirming that the specific geometry and standard support equivalence with the reference gas.
- Use isobutene when a lower-cost decay test is needed and indoor VOC background, ozone, and flammability are controlled.
- Use R134a or similar low-cost platforms for exploratory spatial mapping, not as an automatic replacement for reference-grade analysers.
- Use krypton or other specialized alternatives only where the detection system, regulatory requirements, and validation evidence are available.
- Avoid CO, hydrogen, and flammable hydrocarbons in occupied buildings unless the exposure and ignition risks are explicitly controlled.
- Treat passive VOC tracers as experimental until source stability, humidity response, and analytical recovery are demonstrated under field conditions.

Selection is necessary but not sufficient. A reference-grade tracer will still yield a compromised air-change rate if the mixing, background-stability and calibration assumptions of Section 2.2.2 are unmet, and the sources of error in Section 5.6 act largely independently of which gas was chosen.

5.8 Conclusion

The literature does not support a single universal tracer gas for IAQ research. It supports a hierarchy of fit-for-purpose choices. SF_6 remains the most defensible short-term reference, PFTs remain the strongest passive multizone technology, and CO_2 remains the most practical occupied-building and low-cost tracer. N_2O and krypton can replace SF_6 in selected validated applications, while isobutene and R134a show how lower-cost systems may extend tracer measurements to more locations. Methane, CO, ethylene, radioactive noble gases, hydrogen, helium, and passive VOC candidates are important in the historical or specialized literature but should not be treated as general-purpose solutions.

The most important methodological conclusion is that gas identity is only one part of tracer-gas accuracy. Mixing, sampling position, source stability, time-varying airflow, background concentration, and detector behaviour often dominate the final uncertainty. A defensible IAQ study should therefore report the tracer-selection rationale, release rate, concentration range, detector, calibration, sampling locations, mixing procedure, airflow conditions, background correction, and uncertainty, rather than simply the name of the gas used [3], [2].

Chapter 6

Low-Cost Sensor Technologies for IAQ

Chapter 5 assessed tracer gases as measurement media and judged each against the four criteria established in Section 2.4. This chapter turns to the instruments. Its subject is the class of devices that has made distributed indoor measurement affordable, and its purpose is to establish what those devices can be trusted to report, because Chapter 7 asks them to perform a tracer-gas measurement, and that is a far more demanding task than reporting an indoor concentration.

The chapter is organised by sensing principle rather than by pollutant, because the principle determines the failure mode. A non-dispersive infrared cell and a metal-oxide layer fail in different ways for different reasons, and a chapter arranged pollutant by pollutant obscures the fact that the same physics recurs across targets. Section 6.1 sets out the scope and the maturity gradient that runs through the whole field; Sections 6.2 to 6.5 treat the four principal detection principles; Section 6.6 addresses the system layer; and Section 6.7 draws out what the evidence supports.

One theme recurs and is worth stating at the outset. Across every sensing principle examined here, the reported limitation is rarely the transducer's intrinsic sensitivity, but is instead drift, cross-sensitivity, environmental dependence, unit-to-unit variability, or the absence of a calibration valid under deployment conditions, which is to say that it is the measurement system rather than the sensing element. This is the same conclusion Chapter 7 reaches for tracer platforms, and it is not a coincidence.

6.1 Scope and the Maturity Gradient

Indoor air quality research covers both pollutants and the conditions that shape their behaviour. Common targets include particulate matter, carbon dioxide, carbon

monoxide, nitrogen dioxide, ozone, volatile organic compounds, formaldehyde and radon, while temperature and relative humidity appear in almost every system, partly because they matter for comfort and partly because they modify the response of most low-cost sensing elements [61], [62].

The literature consistently distinguishes reference instruments from low-cost sensors, and the distinction is worth preserving precisely. Reference instruments remain the accuracy benchmark, whereas low-cost sensors offer something different in kind rather than in degree, namely broader spatial coverage, longer monitoring periods, and routine deployment at a cost that permits many units where one reference instrument would previously have been placed [61], [63]. The adoption of low-cost sensing is not a claim that these devices are more accurate, but a claim that a different measurement strategy becomes possible, and that the strategy is worth the accuracy penalty for some questions and not for others.

6.1.1 The maturity gradient

The most important structural feature of this literature is that "low-cost IAQ sensor" is not one technology at a uniform stage of development. Maturity varies sharply by target, and conflating the mature and immature ends of the range is the commonest way the field's readiness is overstated.

Carbon dioxide and particulate matter sit at the mature end. Both have an established physical detection principle, both have been evaluated against reference instruments repeatedly, and both support quantitative use after appropriate calibration [61], [64]. Volatile organic compounds, formaldehyde and the reactive gases sit at the immature end, where selectivity, drift and environmental sensitivity remain unresolved and where the reported quantity is often a proxy rather than a concentration [65], [66].

The distance between these ends is quantified in the same study. In a controlled assessment of low-cost environmental monitors against reference instrumentation, De Manega and colleagues found acceptable agreement for carbon dioxide within 15 per cent, while the same devices produced diverging results with poor quantitative agreement for other pollutants measured simultaneously [67]. The devices were not differently engineered for different targets; the physics of the detection principles differ, and the consequences propagate directly to the data.

An independent systematic review of the same field, published as this work was completed, arrives at a similar position from a much wider search and is worth recording as an external check. Silva and colleagues screened 2736 records against accuracy, application and limitation criteria and retained 24 studies that met their inclusion and quality requirements, and of those 24 only four satisfied the technical and legal requirements for institutional or regulatory use [68]. Their substantive findings match the gradient set out above, in that laser-scattering particulate sensors

achieve acceptable accuracy under controlled conditions but vary considerably in the field, most uncalibrated or factory-calibrated devices show errors above 10 per cent, and nominally identical units of the same model tested simultaneously in one chamber returned coefficients of variation between 12 and 27 per cent depending on particle size and concentration. Two reviews conducted independently, with different search strategies and different inclusion rules, converging on the same maturity ordering is stronger evidence for that ordering than either provides alone.

Table 6.1: Detection principles, targets and dominant failure modes

Principle	**Typical target**	**Dominant failure mode**	**Maturity for quantitative use**
NDIR	CO_2	Baseline drift; pressure and temperature dependence; single-point factory calibration	High after local calibration [61], [59]
Optical scattering	PM_1, $PM_{2.5}$, PM_{10}	Response depends on particle size distribution, composition and refractive index; humidity growth	Moderate to high; source-dependent [64], [69]
Metal-oxide	TVOC proxies, mixed gases	Poor selectivity; irreversible drift; strong humidity dependence	Low for concentration; usable for event detection [65], [70]
Electrochemical	CO, NO_2, O_3, HCHO	Cross-sensitivity between target gases; cell ageing; temperature dependence	Moderate, target-dependent [71], [72]
Photoionisation	VOC mixtures	Responds to total ionisable VOC, not to a species	Moderate; often not low-cost [63]
Temperature and RH	Thermal and moisture state	Enclosure self-heating; poor airflow around the element	Very high [73]

Figure 6.1 places the principal targets on two axes that matter for this thesis, namely how well the quantity can be measured and how directly it supports a ventilation inference. The two are not the same, and the separation explains why carbon dioxide dominates the ventilation literature while particulate matter dominates the exposure literature.

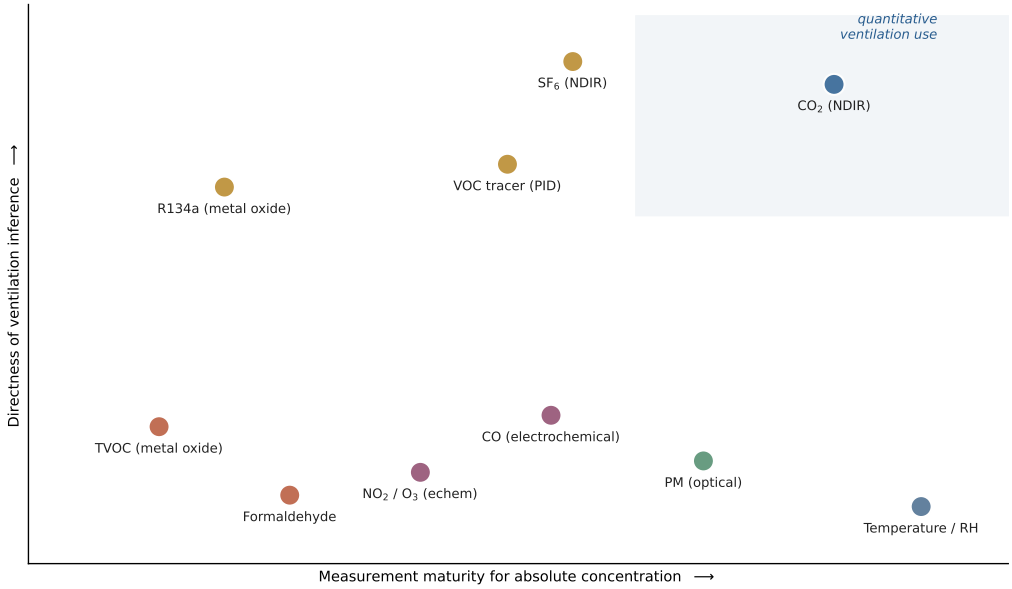


Figure 6.1: Low-cost sensing targets positioned by measurement maturity and by directness of the ventilation inference they support. Only carbon dioxide occupies the upper right, which is why it dominates low-cost ventilation work despite being an imperfect tracer.

6.2 Non-Dispersive Infrared Sensing

Non-dispersive infrared detection is the most mature low-cost sensing principle for indoor use and the only one that this thesis relies on quantitatively. A source illuminates a gas cell, the target gas absorbs at a characteristic wavelength, and the attenuation gives concentration. The principle is specific by construction, because the absorption band is a property of the molecule rather than of the sensing surface, and this specificity is what separates NDIR from the adsorption-based principles discussed in Section 6.4.

6.2.1 Performance and its limits

Review-level evidence consistently identifies NDIR as the most dependable low-cost option for indoor carbon dioxide [61], [59], [74]. Individual evaluations support this with quantitative agreement after calibration, since the Arduino-based device examined in Chapter 7 reached a mean error below 3 per cent against a reference instrument after three rounds of linear calibration [75], and the broader assessment cited above found carbon dioxide agreement within 15 per cent where other targets on the same devices did not agree at all [67].

The qualification is that this performance is conditional on local calibration rather than on factory settings. An on-site evaluation of twenty-three low-cost CO_2 sensors found near-perfect linearity against the reference ($R^2 = 0.999$) but substantial variability in absolute terms, with a mean root-mean-square error of 18.3 ppm before correction, and applying correction factors reduced that error by 27 per cent and reduced out-of-range measurements by 43 per cent [76]. Linearity and accuracy are different properties, and a sensor can be reliably linear while being consistently wrong, which is precisely the condition a single-point factory calibration produces.

A second qualification is that the sensor is often not the binding constraint on the measurement. Marín and colleagues instrumented classrooms with CO_2 modules of differing specification and found that monitor position and the room's ventilation strategy routinely outweighed the differences between sensors, producing spatial discrepancies beyond 100 ppm; field measurements elsewhere have recorded differences of up to 242 ppm between points in the same ventilated room [77]. Against a manufacturer's specification of roughly ±30 to ±50 ppm, a placement-induced discrepancy of 100 ppm or more means that where the module is mounted determines the result more than which module is chosen. The consequence for a network is that spatial sampling and node siting are part of the measurement specification, not deployment details, and Section 7.9 returns to this as a design rule.

6.2.2 What carbon dioxide can and cannot represent

The more consequential limitation is not instrumental. Carbon dioxide concentration is not ventilation rate. It depends on occupancy, activity level, outdoor background and room mixing, and the inference from one to the other carries assumptions that the sensor cannot verify [40], [78], [74].

Sensor position is part of this. Pei and colleagues showed that the location of the CO_2 sensor changes the behaviour of demand-controlled ventilation directly, so that two identically calibrated instruments in the same room can drive different control outcomes [44]. Under good mixing a single sensor gives a usable signal; under poor mixing, displacement ventilation or natural ventilation the same signal may mislead [45], [46]. The limitation therefore sits in the physical model rather than in the transducer, and no improvement in sensor accuracy addresses it. This point returns in Section 7.9 of the following chapter as a design rule.

6.3 Optical Particle Sensing

Optical scattering sensors infer particle mass concentration from the light a particle population scatters as it passes through a beam. The measurement is indirect

in a way NDIR is not, because scattering depends on particle size distribution, composition and refractive index, none of which the sensor measures, so the conversion to mass concentration rests on an assumed aerosol.

This assumption is the dominant source of error, and it is why the same device performs differently against different sources. Evaluations under varying laboratory conditions find that agreement with reference instruments changes with the aerosol being measured rather than remaining a fixed property of the sensor [79], [80]. He and colleagues quantified this by challenging two consumer monitors with several aerosol types and comparing them against a reference photometer under both laboratory and indoor cooking conditions, finding that the regression slope shifted substantially with aerosol type and that the coefficient of variation between three nominally identical units of the same model ran from 15.2 to 21.0 per cent for one device and 12.5 to 28.7 per cent for the other [81]. The calibration is thus specific to an aerosol the deployment does not control. Humidity compounds this, because hygroscopic particles grow in humid air and scatter more light without any change in dry mass.

The error is also range-dependent in a way that a single quoted accuracy conceals. A review of calibration practice for low-cost particulate sensors finds linearity degrading and absolute error rising at elevated concentrations, consistent with counting saturation and coincidence losses, and concludes that performance should be reported by concentration range and source class rather than as one global slope [82]. A normalised error that stays near 20 per cent while the absolute error grows is informative for ranking exposures and misleading if read as mass accuracy.

The scale of unit-to-unit variability was established convincingly by Ahn and colleagues, who evaluated 264 low-cost dust sensors in the laboratory against an exponentially decaying particle concentration and classified them into four groups by accuracy [69]. That a population of nominally comparable sensors separates into four distinct accuracy classes is the central practical finding, since specifying a sensor model does not specify its performance, and a deployment that assumes interchangeability between units is assuming something the evidence contradicts.

The consequence for practice is that optical particle sensors are strong for detecting events, tracking relative change and mapping spatial gradients, and weaker for absolute mass concentration unless calibrated against the aerosol of interest [64]. The distinction matters for exposure assessment, where the absolute value carries the meaning, more than for source detection, where the change does.

6.4 Metal-Oxide Sensing and the Selectivity Problem

Metal-oxide semiconductor sensors detect gases through changes in the conductivity of a heated oxide layer as molecules adsorb onto its surface. The mechanism is sensitive and inexpensive, and it is fundamentally non-selective, because the layer responds to whatever adsorbs, so the output is a function of the whole gas mixture rather than of a named species.

6.4.1 Why the output is a proxy

This is the reason the reported quantity in most low-cost VOC sensing is a total volatile organic compound index rather than a concentration. The reading is a weighted response to a mixture whose composition is unknown, and two rooms with identical index values may have entirely different chemistry [65], [66]. Presenting such a reading as a concentration in micrograms per cubic metre attributes a specificity the measurement does not possess.

Kolarik and colleagues demonstrated the composition dependence rather than asserting it, exposing several commercial metal-oxide VOC sensors to distinct emission scenarios in a test room while a proton-transfer-reaction time-of-flight mass spectrometer identified which compounds were actually present [83]. The mixtures differed sharply between scenarios, with organic acids contributing about half the total VOC concentration under human bioeffluent emission while a wholly different set of compounds dominated a material-emission scenario. The sensors responded to both, but not in a fixed proportion to the measured total, which is the behaviour the mechanism predicts and the reason a single calibration factor cannot be transported between source types. The same study is nonetheless the strongest available case for using these sensors as a ventilation signal, because comparing a VOC sensor with a CO_2 sensor in an office room showed that the two agreed ventilation was required for 49 per cent of occupied time, and that for a further 11 per cent the VOC sensor alone called for it, detecting a demand of non-occupant origin to which a CO_2-based controller is blind by construction.

Drift compounds the problem. The sensing layer changes irreversibly during operation, so the response to a fixed gas mixture is not stable over months. In a field study of metal-oxide sensors, Baur and colleagues documented this drift directly and evaluated calibration strategies against it, finding that a model trained on an initial calibration degrades as the sensor ages and that periodic recalibration or explicit drift compensation is required to maintain usable output [70]. Humidity dependence adds a further term, since water vapour competes for the same adsorption sites as the target gases.

6.4.2 Where these sensors are nonetheless useful

None of this makes metal-oxide sensing worthless, and the mistake in the opposite direction is as common. These devices are genuinely good at detecting that something has changed. Event detection, source identification by temporal correlation, and relative spatial mapping are all well served by a sensitive, non-selective, inexpensive detector, and Chapter 7 shows that a metal-oxide element can map a deliberately released refrigerant tracer densely enough to reveal dispersion pathways [50]. The requirement is that the claim made of the data matches what the principle supports.

6.5 Electrochemical and Photoionisation Detection

Electrochemical cells detect a target gas through the current generated by its oxidation or reduction at an electrode. They are the standard low-cost route for carbon monoxide, nitrogen dioxide and ozone, and are used for formaldehyde [72], [71]. Selectivity is better than for metal-oxide devices because the electrode chemistry is targeted, but it is not absolute, since cells respond to other electroactive species, so cross-sensitivity between the very gases a multi-pollutant device measures simultaneously is a recognised difficulty.

Formaldehyde illustrates the compounded difficulty well. It is a priority indoor pollutant with exposure limits that differ between countries, and its emission from building materials is strongly temperature dependent, with Chattopadhyay and colleagues noting emissions doubling for roughly every 7 °C increase [71]. A sensor must therefore separate a genuine change in source strength from a change driven by room temperature, while itself being temperature sensitive, and a calibration that does not span the deployment temperature range cannot support that separation.

Electrochemical cells also age, since the electrolyte is consumed and sensitivity declines over a service life measured in months to a small number of years. A deployment intended to run for a year therefore needs a recalibration plan rather than merely an initial calibration.

Photoionisation detectors occupy a different position. A lamp ionises molecules below its photon energy and the resulting current is measured, giving a fast, sensitive response to total ionisable volatile organic compounds. The response is still to a class rather than a species, but the class is better defined than the metal-oxide response and the device is considerably more stable. The cost sits between metal-oxide and reference instrumentation, and Chapter 7 shows this to be the decisive advantage where a deliberately released VOC tracer must be distinguished from indoor background [49].

6.6 The System Layer

A large part of the recent literature treats indoor air quality sensing as a complete digital system rather than a set of transducers, and the shift is consequential because most reported failures occur above the sensor.

Arduino, Raspberry Pi, ESP8266 and ESP32 platforms recur because they are inexpensive, flexible and easily connected to network services [73], [84], [85]. Wi-Fi is the most common link, with Bluetooth, ZigBee and lower-power alternatives appearing where range or energy budget dominates [84], [63]. Dashboards, cloud storage and mobile interfaces are now standard [84], [86].

The consistent finding across system-level studies is that this maturity is uneven in a specific way, in that architecture is more developed than validation. It is easier to publish a working networked deployment than a defensible accuracy claim, and the literature reflects that asymmetry [87], [88].

Two system-level results deserve particular note because they quantify what deployment costs in reliability.

The first is unit-to-unit variability and outright failure. In a co-location study of 87 low-cost environmental monitors, Manibusan and colleagues found that 5 of the 87 units failed outright during the exercise [89]. A failure rate near six per cent is manageable in a network designed for it and fatal in a deployment that assumes every node reports. Dense low-cost networks acquire a reliability engineering problem that single-instrument measurement does not have.

The second is that co-location is now the accepted route to a defensible claim, and the reporting standard has become explicit. Zaky and colleagues evaluated consumer-grade CO_2 and $PM_{2.5}$ monitors against two reference instruments using time-series analysis, linear regression, Pearson correlation, root-mean-square error, Bland–Altman limits of agreement and paired testing, and used the combination to separate precision from accuracy, finding one device with higher correlation and narrower limits of agreement but larger bias than another, so that it tracked relative change better while reporting absolute values less well [90]. That separation is the analytically important one, and it is invisible to any single accuracy figure.

Laboratory protocols now exist to support this at scale, both for controlled decaying concentrations [69] and for structured performance assessment under laboratory and simulated-real conditions [91], [92]. The instruments for validating low-cost sensors have matured faster than the practice of using them.

6.7 What the Evidence Supports

Three conclusions follow from this chapter and carry into the next.

Maturity is target-specific, and the phrase "low-cost IAQ sensor"

should not be used without one. Carbon dioxide by NDIR and particulate matter by optical scattering support quantitative use after appropriate calibration. Metal-oxide VOC sensing supports event detection and relative comparison. Treating these as a single category, as the review literature sometimes does, produces claims that are true of the mature end and false of the other.

The limiting factor is almost never the transducer. Every principle examined here reports its dominant error as drift, cross-sensitivity, environmental dependence, unit-to-unit variability, or the absence of a calibration valid under deployment conditions. These are properties of the measurement system, and they are addressed by calibration design, placement, recalibration schedules and network architecture rather than by a better sensing element.

Calibration and placement are not preparatory steps but part of the measurement. A sensor calibrated at one point, in one temperature and humidity condition, against one aerosol, positioned without regard to room airflow, does not have a knowable accuracy in deployment, whatever its datasheet says. The studies in this chapter that report credible agreement all did the calibration work, and the studies that report position effects show that the same instrument gives different answers from different places in the same room.

These conclusions set the terms for Chapter 7. A tracer-gas measurement demands more than an indoor concentration reading, because it requires that the concentration be attributable to a known release, resolved over a decay that traverses the low end of the sensor's range, and interpreted through a mixing assumption the sensor cannot check. The sensing principles reviewed here are the raw material for that task, and the gap between what they can measure and what a tracer measurement requires is the subject of the next chapter.

Chapter 7

Low-Cost Tracer-Gas Sensing Systems

Chapter 6 examined low-cost sensors as transducers, and Chapter 5 examined tracer gases as measurement media. This chapter takes the intersection of the two, namely instruments that attempt to perform a tracer-gas measurement with inexpensive hardware. Section 4.3 established that this intersection is where the corpus is thinnest, and the chapter is written with that scarcity in view. It does not survey a mature field, because there is no mature field to survey, and it instead examines a small number of prototype platforms closely enough to say what each has actually demonstrated, separating that from what remains assumed.

The organising question is not whether an inexpensive sensor can detect a tracer gas, since several can. The question is whether such a sensor can support the inference that the tracer-gas method exists to make, whether a ventilation rate, an interzonal flow or a dispersion pathway, to a standard that would let it stand in for a reference instrument. That is a considerably stronger claim, and the distance between the two is the subject of this chapter.

7.1 From Classical Methods to Low-Cost Platforms

The two literatures described in Section 4.1 approach the same measurement from opposite directions, and the difference in starting point explains most of what follows.

The classical tracer-gas tradition began from the physics. It derived the single-zone mass balance, established the decay, constant-injection and constant-concentration designs, and specified the instrumentation needed to satisfy the

assumptions those designs rest on [5], [4]. Detection was expensive because the method demanded it, requiring electron-capture chromatography for sulfur hexafluoride, thermal desorption and gas chromatography for perfluorocarbons, and photoacoustic spectroscopy where continuous multi-point sampling was needed. Cost was a consequence of the accuracy target rather than a design variable.

The low-cost sensing tradition began from the hardware. Inexpensive transducers, microcontrollers and wireless modules made dense, continuous, distributed measurement possible for the first time, and the literature that followed asked what could be done with that capability [61], [63]. Ventilation was one application among many, and in most of this work the measured quantity is a concentration rather than a flow.

A low-cost tracer-gas platform has to satisfy both traditions at once, and the two impose different requirements. Table 7.1 sets out the pathways by which ventilation is inferred in the retrieved literature, and makes clear that they are not interchangeable.

Table 7.1: Pathways to a ventilation estimate, and requirements

Pathway	**Typical sensing**	**What must hold**	**Cost driver**
Deliberate tracer, controlled release	GC-ECD, PAS, NDIR, PID, MOS	Known source rate; conserved tracer; adequate mixing	Detector sensitivity and selectivity
Occupant-generated CO_2	Low-cost NDIR	Known generation rate; stable outdoor background; adequate mixing	Sensor count and placement
Passive integrated sampling	PFT sources and adsorption tubes	Stable emission; laboratory analysis; long averaging	Laboratory analysis, not field hardware
Envelope pressurisation	Fan and pressure transducer	Steady induced pressure; leakage model	Fan and rig, not sensing

The distinction that matters for this chapter is between the first row and the second. A deliberate tracer gives control of the source term at the cost of requiring a detector specific enough to see that tracer against the indoor background, whereas occupant-generated CO_2 removes the release equipment entirely but surrenders control of the source strength, which then has to be estimated [40], [2]. Low-cost hardware has made real progress on the second pathway and much less on the first, and the reason is not economic but analytical, since cheap detectors are rarely selective and a tracer measurement without selectivity is a measurement of something else.

7.2 The Evidence Boundary

Before examining individual platforms it is necessary to state plainly how much evidence exists, because the volume of publication on low-cost IAQ sensing gives a misleading impression of the volume of evidence on low-cost *tracer* sensing.

Within the 226-record analysis corpus, 98 records address tracer-gas methods and 74 address low-cost sensing, but only nine address both, three of those name the detection principle they used, and two report a comparison against a reference instrument (Section 4.3). The platforms discussed in the remainder of this chapter are therefore not a sample of a population. They are close to the population.

One distinction governs everything that follows, and it is worth making before any hardware is described. A system tested in a chamber or a controlled laboratory layout has demonstrated that its detector responds correctly to a tracer under conditions the experimenter arranged, whereas a system tested in an occupied building has demonstrated something considerably harder, namely that the response survives real mixing, real occupancy, real background variation and real airflow that changes while the measurement is running. The retrieved platforms divide unevenly across this boundary, and the division does not follow cost, sophistication or recency. The two most technically developed systems in the set, the R134a wireless grid and the TEBE-SENSE SF_6 platform, are both laboratory demonstrations, while the systems with genuine occupied-building results are the simpler CO_2 devices.

7.3 Engineering Comparison of Retrieved Platforms

Table 7.2 compares the six platforms in the corpus that attempt a tracer-gas measurement with low-cost hardware. It is deliberately specific, recording hardware, working range, time resolution, the validation actually performed, and the setting in which it was performed. Vague comparison at the level of "low-cost systems show promise" is precisely what this chapter is intended to avoid.

Table 7.2: Low-cost tracer-sensing platforms in the corpus

Tracer and system	Detector and platform	Range and time resolution	Validation performed	Setting
R134a wireless grid [50]	Figaro TGS 3830 tin-dioxide sensor; NuMicro microcontroller; ESP8266 Wi-Fi to cloud; passive diffusion	5–100 ppm; 10 s sampling cycle	ACH compared against an SF_6 reference system: mean error 8%, range 1.9–14.4%	Controlled laboratory layout
SF_6 TEBE-SENSE [54]	NDIR detector; SAMD21 Cortex-M0; SHT85 temperature and humidity; internal fan; Bluetooth Low Energy	500 ppm (two units) and 2000 ppm full scale; internal sampling to 50 Hz	Decay-derived ventilation rate within ±20% of reference for 1–7 h^{-1}; higher-range unit dispersed more	40 L test box
Isobutene PID [49]	RAE PGM-7240 photoionisation detector, 10.6 eV lamp; zeroed on clean air, calibrated on 10 ppm isobutene	Tracer raised to ≈10–100 ppm to clear VOC background	Air-change rate agreed with a simultaneous SF_6 photoacoustic measurement at ≈1.0 h^{-1}; roughly one-fifth the cost	Occupied apartment buildings
CO_2 classroom mesh [93]	Five Cozir Ambient NDIR nodes; wireless mesh; three pressure and temperature sensors	Room scale; decay recorded after occupancy	Internal: node-to-node spatial uniformity check plus a standard decay protocol; no independent tracer reference	Occupied 383 m^3 classroom
CO_2 school network [94]	Miniature photoacoustic NDIR: modulated mid-infrared LED, MEMS microphone, PSoC processor; Z-Wave to web platform	Calibrated on nine certified concentrations, 300–700 ppm, extended upward	Certified-gas calibration and operational plausibility under real ventilation actions	Occupied school, long term

Continued on next page...

Table 7.2: Low-cost tracer-sensing platforms in the corpus (continued)

Tracer and system	Detector and platform	Range and time resolution	Validation performed	Setting
CO_2 Arduino device [75]	DFRobot SEN0219 NDIR on ELEGOO Mega 2560; placement chosen with CFD	5 min sampling interval	Three rounds of linear calibration against a KIMO HQ-210 reference: mean error below 3%, maximum below 8%, $R^2 = 0.9931$	Occupied rooms, several ventilation modes

Three observations follow from the table, and they structure the rest of the chapter.

The first is that the tracer determines the detector, and the detector determines the cost floor. The only platforms using a deliberate, controllable tracer are the three that invest in a selective detector, being a photoionisation detector for isobutene, an NDIR cell tuned for SF_6, and a tin-dioxide sensor for R134a. The CO_2 systems avoid the problem by using a tracer the occupants supply, and pay for it in source-term uncertainty.

The second is that time resolution and application are tightly coupled. Ten-second sampling supports transient dispersion mapping; five-minute sampling does not, and is suited to slow decay and ventilation trends. This is a design decision rather than a quality difference, but it is frequently reported as though it were incidental.

The third is the pattern already noted in Section 7.2, which is that the two systems comparing themselves against a reference tracer instrument under controlled conditions are laboratory studies, while the systems that operate in occupied buildings mostly validate themselves internally. Only one platform in the table does both, and it is the oldest.

7.4 Semiconductor Sensing of Refrigerant Tracers

The R134a wireless grid of Tsang and colleagues is the clearest demonstration in the corpus of what dense, inexpensive spatial sampling can contribute to a tracer measurement [50]. Each node combines a Figaro TGS 3830 tin-dioxide sensor with a microcontroller and an ESP8266 Wi-Fi module, transmitting directly to a cloud database. Sensing is by passive diffusion rather than a pump, which lowers cost and power draw at the price of response time.

The performance figures are specific enough to assess. The stated detection range is 5–100 ppm with a ten-second sampling cycle, and air-change rates derived from the grid agreed with those from an SF_6 reference system to a mean error of 8%, with individual comparisons spanning 1.9 to 14.4 per cent. For a sensor of this class that is a good result, and the ten-second cycle is fast enough to resolve dispersion behaviour that a five-minute instrument would average away entirely.

The qualifications are equally specific. The work was conducted in a controlled laboratory layout rather than an occupied building, and the properties that most threaten a semiconductor sensor in the field were consequently not tested, among them long-term baseline drift, interaction with building materials, cross-sensitivity to the mixed and variable VOC background of a real interior, and packet loss when wireless conditions are poor. Performance also degraded outside the stated operating range, which matters because a decay experiment necessarily traverses the low-concentration end of the range as the tracer disappears, which is precisely where a sensor calibrated at higher concentrations is least trustworthy and precisely where the exponential fit is most sensitive to error.

The constructive reading is that this platform demonstrates a two-tier architecture rather than a replacement instrument. Inexpensive nodes map the relative concentration field with a spatial density no reference instrument could economically match; at least one reference node or analyser supplies absolute concentration and detects network drift. A field-ready version would add local buffering so that a network outage does not destroy the run, synchronised clocks so that nodes started at different times remain comparable, co-logged temperature and humidity, and a calibration valid across the full range the experiment will actually traverse.

7.5 Occupant-Generated CO_2 in Real Buildings

Carbon dioxide is the best-supported low-cost tracer for occupied buildings, and the reason is logistical rather than analytical. It requires no cylinder, no release equipment and no permission to introduce a foreign substance into a space where people are working; the occupants generate it, and compact infrared sensors measure it well. The corresponding weakness is that the source term is uncontrolled and the background varies, so the systems below are properly understood as low-cost ventilation instruments rather than as tracer-gas instruments in the classical sense [40].

Three field systems in the corpus illustrate three different levels of ambition.

The classroom mesh of Benedettelli and colleagues deployed five low-power Cozir Ambient NDIR nodes with three pressure and temperature sensors across a 383 m^3 classroom, with three nodes inside the room, one in the hallway and one outdoors [93]. Occupants raised the concentration during a lecture, after which the room was

emptied and sealed and the decay recorded. The design is instructive because it combines distributed sensing, an explicit decay protocol, a spatial uniformity check and a measurement of natural leakage in a single occupied-building experiment. Its limitation is that validation is internal, since the study establishes that the nodes agree with one another well enough to justify a single-zone decay fit, which is a necessary condition for the analysis, but it does not establish that the resulting air-change rate is correct, because no independent tracer or calibrated airflow reference was run alongside.

The school network of Pérez and colleagues is a more complete platform [94]. Its node uses miniature photoacoustic NDIR sensing, combining a modulated mid-infrared LED, a CO_2-filled cell and a MEMS microphone, with a PSoC processor, integrated temperature and humidity sensing and Z-Wave communication to a web platform. Calibration used nine certified concentrations between 300 and 700 ppm, extended upward. The field results are operationally valuable, since short breaks between lessons did not restore background concentrations whereas deliberate window-and-door ventilation did, and that is exactly the kind of finding a building manager can act on. It is not, however, equivalent to a controlled tracer test, and the study does not claim otherwise.

The Arduino-based device of Pastor-Fernández and colleagues is the simplest of the three and the most carefully calibrated [75]. A DFRobot SEN0219 NDIR sensor on an ELEGOO Mega 2560 was calibrated in three rounds against a KIMO HQ-210 reference, reducing mean error below 3 per cent with a maximum below 8 per cent and a coefficient of determination of 0.9931. Notably, computational fluid dynamics was used to choose sensor positions that avoided vortices and stagnant zones. The five-minute sampling interval rules out transient dispersion work, and no full independent air-change comparison is reported for each ventilation scenario, so the device is best suited to screening, control and slow decay. Its methodological contribution is the demonstration that placement deserves the same analytical attention as calibration, a point developed further in Section 7.8.

A fourth contribution approaches the same pathway from the application side rather than the hardware side. Singh and Dewan present a low-cost tool that uses carbon dioxide as a biomarker of rebreathed air in order to derive the air-change rate needed for airborne infection control, with the sensor supplied with ambient room air so that the reading is representative [95]. The interest here is not the instrument, which is conventional, but the framing, since the measurement is specified backwards from a required ventilation rate rather than forwards from a sensor specification. That inversion is uncommon in this literature, and it makes the required measurement uncertainty a design input rather than a reported outcome.

Taken together these systems show that the CO_2 pathway is genuinely usable in occupied buildings, and equally that its validation ceiling is set by the source term rather than by the sensor. Improving the detector does not resolve uncertainty in

how much CO_2 the occupants produced, nor in what the outdoor background was at the time [2], [59].

7.6 Photoionisation Detection of a Deliberate VOC Tracer

The isobutene and photoionisation-detector combination reported by Hori and Mizoguchi is the strongest lower-cost alternative in the corpus that uses a deliberate, controllable tracer in an occupied building [49]. A RAE PGM-7240 with a 10.6 eV lamp was zeroed on clean air and calibrated with 10 ppm isobutene. Because a photoionisation detector responds to the general VOC background as well as to the chosen tracer, the working concentration was raised to roughly 10–100 ppm, several times the ambient VOC level, so that the tracer signal could be separated from it.

The comparison is the part that matters. In apartment buildings with both mechanical and passive ventilation, the isobutene system and a simultaneous SF_6 photoacoustic measurement both returned an air-change rate of approximately $1.0\,h^{-1}$, at roughly one-fifth of the cost of the photoacoustic system. This is one of only two reference-instrument comparisons in the corpus conducted outside a laboratory, and it is the only one in which a deliberate non-CO_2 tracer was used in an occupied building.

The trade-offs are real. A photoionisation detector costs more than a metal-oxide sensor, retains some sensitivity to background VOCs even at elevated tracer concentrations, and requires enough tracer to be released to establish that margin. The method therefore suits short tests in spaces that can be temporarily controlled, and suits neither continuous occupied monitoring nor unattended dense grids. Within that envelope it is the most convincing demonstration available that a substantially cheaper detector can reproduce a reference tracer result in a real building.

7.7 NDIR Detection of SF_6 in the TEBE-SENSE Platform

The TEBE-SENSE system is the most developed low-cost SF_6 platform in the retrieved literature and the one most directly relevant to the experimental direction this review points towards [54]. It pairs an NDIR detector with a SAMD21 Cortex-M0 microcontroller, SHT85 temperature and humidity sensing, a small internal fan and Bluetooth Low Energy communication. Signals are digitised at 12-bit resolution with an internal sampling frequency of 50 Hz, and tracer-decay data are fitted by least squares.

Three findings deserve emphasis. First, the system estimated ventilation rates within ±20 per cent of the reference for rates between 1 and 7 h^{-1} in a 40 L test box, which establishes that NDIR detection of SF_6 at low-cost is analytically feasible rather than merely conceivable. Second, full-scale selection proved consequential, since modules were tested at 500 ppm and 2000 ppm full scale and the higher-range module showed markedly greater dispersion, overestimating flow rates at low concentrations. This is an instructive result, because it shows the error entering through instrument configuration rather than through the sensing principle, in that a 2000 ppm module used on a decay that ends near the bottom of its range is being asked to resolve a difference it cannot resolve. Third, the internal fan improves flushing and shortens response time, but introduces a design tension examined in Section 7.8, because forced flow through the enclosure alters the local flow field the sensor is meant to sample.

The evidence boundary applies with full force here. The demonstration is laboratory-based, in a 40 L box rather than a room, and the study does not establish operation in a real building, long-term drift, cross-sensitivity in mixed indoor air, or behaviour under complex interzonal flows. The platform should be read as a prototype architecture with a credible analytical basis rather than as a field-validated replacement for photoacoustic spectroscopy. That is a strong position for a prototype to occupy, and the gap between it and a field instrument is a well-defined engineering programme rather than an open research question, which is precisely why it is the natural point of departure for experimental work.

7.8 What "Validation" Means Across These Systems

The retrieved platforms all describe themselves as validated, and they mean at least five different things by it. Conflating them is the single most common way the low-cost tracer literature overstates its own maturity, and separating them is the most useful analytical contribution this chapter can make.

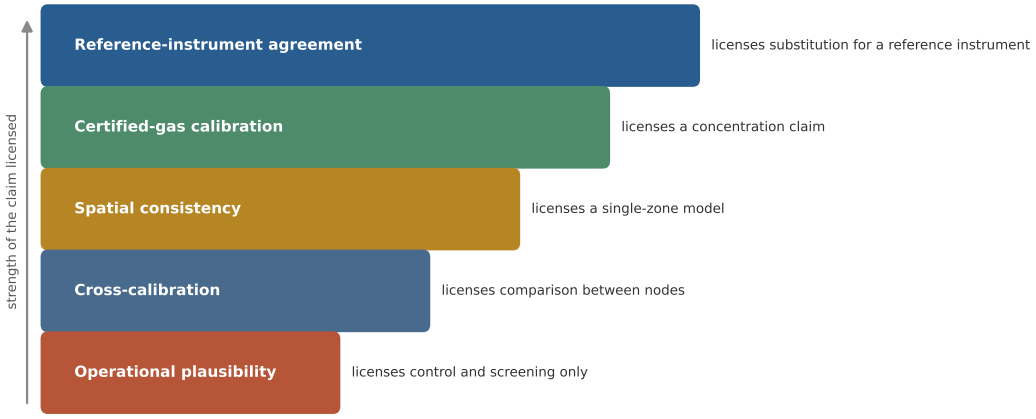


Figure 7.1: The five senses of validation ordered by the strength of the claim each licenses, with the claim licensed by each rung given alongside it. Only the top rung tests the quantity a tracer measurement exists to produce, and a platform can satisfy every rung below it while still reporting a biased ventilation rate, because concentration accuracy is compatible with error in the source term, in the background estimate or in the mixing assumption. Table 7.3 gives the corresponding evidence and its limits.

Table 7.3: Five senses of validation in the low-cost tracer literature

Sense	What is demonstrated	What is not
Reference-instrument agreement	The system reproduces a result from an accepted instrument on the same event [50], [49], [54]	Nothing beyond the conditions tested; two of the three were laboratory studies
Certified-gas calibration	Concentration response is correct against known standards [94], [75]	That the inferred ventilation rate is correct at room scale
Spatial consistency	Nodes agree closely enough to justify a single-zone model [93]	Absolute accuracy; all nodes may share a bias
Operational plausibility	Response follows expected physical behaviour under occupancy and window opening [94], [75]	Any quantitative claim; useful for control and screening only
Cross-calibration	Inter-node variability is reduced by co-location against a network-derived reference [96]	Absolute accuracy of the network as a whole

The hierarchy is not arbitrary. Only the first sense tests the quantity the tracer method exists to produce; the remaining four test necessary conditions. A system may satisfy all four of the lower senses and still return the wrong air-change rate, because concentration accuracy, spatial uniformity, plausible behaviour and inter-node consistency are all compatible with a systematic error in the source term, the background estimate or the mixing assumption.

This is why the corpus statistic reported in Section 4.3, of two records discussing reference validation, is more serious than it first appears. It does not mean that only two systems were tested, but that only two tested the thing that would license substituting a cheap instrument for an expensive one.

The corresponding recommendation for reporting is straightforward and rarely followed, which is that concentration error, decay-fit uncertainty, spatial variability across nodes, and agreement with a reference method should be reported as four separate quantities. A single accuracy figure conceals which of them is limiting, and it is usually not the sensor. Recent co-location work outside the tracer setting shows what the reporting standard can look like when it is taken seriously, using regression, correlation, root-mean-square error, Bland–Altman limits of agreement and paired testing to separate precision from accuracy [90], and laboratory protocols now exist for evaluating large numbers of low-cost units against a controlled decaying concentration [69], [91], [92].

7.9 Design Guidance for a Low-Cost Tracer Platform

The platforms reviewed above fail and succeed in patterned ways, and the pattern supports a set of design rules. They are stated as system-level requirements rather than component choices, because the recurring lesson of Sections 7.4 to 7.7 is that the sensor is rarely the limiting element.

Match the detector to the tracer, not to the budget. NDIR is the strongest low-cost route for CO_2 and is now demonstrated for SF_6. A photoionisation detector is preferable to a generic metal-oxide sensor whenever a VOC tracer must be separated from indoor background. Metal-oxide devices can support inexpensive refrigerant mapping, but their range, drift and cross-sensitivity must be characterised before the data are interpreted as concentrations [50], [49], [54].

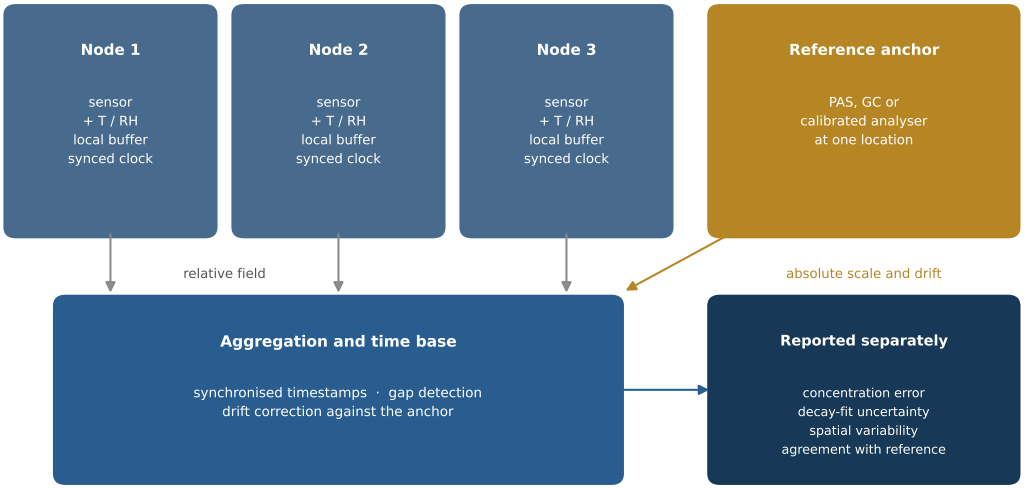


Figure 7.2: The architecture the evidence supports. Inexpensive nodes supply spatial coverage and a relative concentration field, while a single reference instrument supplies absolute scale and detects network drift, and the aggregation layer applies a synchronised time base before error is reported as separate quantities. Local buffering and clock synchronisation are required because direct-to-cloud architectures lose data when connectivity fails and cannot align nodes started at different times, which are the failure modes reported in Section 7.4.

Anchor the grid with a reference. A single photoacoustic analyser, chromatograph or calibrated instrument at one location supplies absolute concentration and reveals network drift, while inexpensive nodes supply spatial coverage, as in Figure 7.2. This is more defensible than treating every node as an independent reference, and it is the architecture the R134a results implicitly recommend.

Calibrate across the range the experiment will traverse. A decay experiment ends at low concentration by construction. Calibration at a single mid-range point leaves the most influential part of the fit unconstrained, and the TEBE-SENSE full-scale result shows how large the resulting error can be [54]. Calibration should cover the working range and include zero and background checks, temperature and humidity conditions, response and recovery time, and sensor-to-sensor variability.

Buffer locally and synchronise clocks. Direct-to-cloud architectures are convenient and fragile. Data were lost when wireless connectivity failed, and nodes started at different times required manual alignment [50]. Each node should write raw data locally and carry a synchronised timestamp before anything is transmitted.

Choose the sampling interval from the physics, not the hardware. Ten seconds suits room-scale dispersion; five minutes suits ventilation trends and slow

decay. Reporting the interval without justifying it against the timescale of the process being measured leaves the reader unable to judge whether the sampling was adequate.

Treat placement as part of the calibration. Height, distance from occupants, proximity to supply diffusers, corners and stagnant zones can dominate the result, and sensor position has been shown to change the outcome of CO_2-based ventilation control directly [44]. Section 6.2 puts a number on the effect, since position and ventilation strategy together produced spatial discrepancies beyond 100 ppm, exceeding the difference between sensor models [77]. Computational fluid dynamics, smoke visualisation or a preliminary multi-point uniformity check should therefore inform position before the measurement campaign begins [75], [93].

Where a single node must represent a room, the position can be specified rather than left to judgement. Ackley and colleagues tested this in three naturally ventilated classrooms, comparing sensors distributed across a horizontal plane at the room centre against single sensors mounted on each of the four walls, and found the variability between the two arrangements to be largely below 100 ppm [97]. Their resulting specification, of head height at approximately 1.5 m, on a vertical wall, away from windows, doors and supply openings, and not directly within the breathing zone of an occupant, is a concrete default for a single-node deployment, and the fact that it was arrived at by measurement rather than convention is what makes it usable.

Size the network from the control objective, not the room count. Node count is normally set by budget or by one sensor per room. Hobson and colleagues addressed it as a design variable, simulating CO_2-based demand-controlled ventilation across a 26-zone office floor under ten occupancy schedules and asking how few sensors would suffice if unsensed zones were controlled from adjacent sensed ones [98]. Sensors in approximately 31 per cent of zones, being eight of twenty-six, enabled control across the whole floor plate with negligible effect on the hours of elevated CO_2, and delivered 7.3 to 17.4 per cent higher median energy savings than controlling each zone individually with the same number of sensors, whereas matching those savings under individual control would have required two to two-and-a-half times the sensing infrastructure. The transferable result is not the 31 per cent, which is specific to that floor plate, but the finding that coverage and control topology trade against each other, so a network should be sized against what is to be inferred and from where.

The same study records a gap between what standards assume and what hardware delivers that bears directly on the calibration rules above. ASHRAE Standard 189.1 and California Title 24 require precision of ± 50 and ± 75 ppm respectively for sensors used in demand-controlled ventilation, and require that recalibration not be needed for at least five years, yet among the units surveyed one batch contained a sensor biased by as much as 500 ppm [98]. A specification

written as a purchasing requirement is not a substitute for verifying the units actually installed.

Manage the fan trade-off explicitly. Forced flow shortens response time and improves flushing, but an enclosure that pushes air past the sensing element creates a local flow field that is not the room's. The choice should be stated and justified rather than left implicit in the mechanical design [54].

7.10 Technology Selection by Application

The evidence supports different choices for different measurement problems, and Table 7.4 states them with the confidence the evidence actually warrants.

Table 7.4: Recommended technology by application, with supporting evidence

Application	Recommended route	Strength of the evidence
Occupied schools, offices and homes; continuous monitoring	Low-cost NDIR CO_2 nodes with certified-gas calibration and considered placement	Good. Multiple occupied-building demonstrations, though validation is calibration and plausibility rather than reference comparison [94], [75], [93]
Short controlled tests needing a deliberate tracer in a real building	Isobutene with a photoionisation detector	Moderate. One occupied-building reference comparison, at roughly one-fifth the cost of photoacoustic [49]
Spatial and transient dispersion mapping	Dense inexpensive nodes with a reference anchor; refrigerant tracer with metal-oxide sensing	Provisional. Strong laboratory result; field behaviour, drift and material interaction untested [50]
Fast decay measurement with a classical tracer	Low-cost NDIR for SF_6	Provisional. Laboratory only; range selection is critical and must match the expected concentration [54]
Long-term average air-change rate	Passive perfluorocarbon sampling with laboratory analysis	Strong but not low-cost in the sense used here: the cost sits in the analysis, not the hardware [33], [99]

Two entries deserve comment. The passive perfluorocarbon route is included because it is the correct answer for long-term averaging and because it exposes an ambiguity in the phrase "low-cost", in that the field hardware is inexpensive and the laboratory analysis is not, so the method is cheap to deploy and expensive to complete. Any cost comparison that counts only the deployed hardware will

therefore mislead. Work on passive tracers is, however, addressing precisely this cost structure. Paralovo and colleagues developed a passive source of deuterated decane chosen so that the tracer can be co-captured and co-analysed alongside the volatile organic compounds a commercial passive indoor-air sampler already collects, with a source design giving stable and repeatable emission and an exponential relation between emission rate and room temperature [58], [99]. If the tracer can ride on an analysis a building study is performing anyway, the marginal cost of the air-change measurement falls sharply, and the distinction between cheap hardware and expensive analysis begins to close from the analysis side.

The second is that no entry in the table is supported by a reference-instrument comparison conducted in an occupied building, with the single exception of the isobutene row. That is the state of the evidence, and it should be read as a description of where work is needed rather than as a reason to avoid low-cost methods.

7.11 Conclusion

Low-cost tracer-gas sensing is at the stage where the analytical feasibility of each pathway has been demonstrated and the field readiness of none has. Inexpensive NDIR can detect SF_6 and CO_2 with usable accuracy; a photoionisation detector can follow a deliberate VOC tracer against indoor background; a semiconductor sensor can map a refrigerant tracer densely enough to reveal dispersion pathways that no affordable reference instrument could resolve. Each of these is a real result.

What has not been demonstrated is that any of these platforms can replace a reference instrument in an occupied building over a sustained period. The two systems with the strongest reference comparisons were tested in a laboratory layout and a 40-litre box respectively; the systems with genuine occupied-building records validate themselves largely against internal consistency and physical plausibility. The one exception, the isobutene comparison, is nearly two decades old and used a detector that is cheap only relative to photoacoustic spectroscopy.

The gap is therefore specific rather than general, and that specificity is useful. It is not that low-cost tracer sensing is unproven in principle, but that the particular experiment which would prove it, namely a low-cost platform and a reference instrument measuring the same decay in an occupied building, repeatedly and over long enough to expose drift, has not been reported for any of the platforms in this corpus. The diagnostic problems such a platform would be applied to, and the experimental programme required to close this gap, follow in the remaining chapters.

Chapter 8

Applications and Diagnostic Use Cases

Chapters 5 to 7 assessed tracer gases, sensing instruments and the platforms combining them as measurement technologies. This chapter addresses what those technologies are used for. Its subject is the diagnostic problem rather than the instrument, and its purpose is to establish which method answers which question, because the literature frequently reports results from one problem as though they answered another.

The chapter is organised around a distinction developed in Section 8.1, which is that the questions usually grouped under "ventilation measurement" are four separate physical problems, measured at different pressure regimes, over different timescales, and against different reference methods. Sections 8.3 to 8.6 treat each in turn, and Section 8.7 states which method is appropriate for which diagnostic goal.

8.1 Four Distinct Measurement Problems

Tracer gases and sensor networks are applied to four problems that share a vocabulary and little else. Collapsing them into a single air-change measurement is the most common structural error in this literature, because a method that is adequate for one may be inapplicable to another.

Table 8.1: The four measurement problems and their reference methods

Problem	Physical quantity	Pressure regime	Strong reference method
Envelope leakage	Leakage area or permeability of the fabric	Imposed, typically 50 Pa	Fan pressurisation [100], [101]
Natural infiltration	Actual outdoor-air exchange in the operating state	Natural, of order 1–4 Pa	Tracer decay or constant concentration [2], [14]
Distribution-system flow	Volumetric flow and leakage within ductwork	Fan-driven, system dependent	Tracer dosing with downstream sampling [102]
Source localisation	Position, strength and timing of an unknown release	Whatever the building is in	No accepted reference; model-based inference [103], [104], [105]

Two consequences follow. The first is that no single method measures all four well. Fan pressurisation gives a repeatable envelope benchmark at an imposed pressure; tracer methods measure actual transport in a selected operating state; sensor networks add spatial and temporal coverage without removing the need for a model connecting concentration to flow.

The second is a distinction that governs the interpretation of every result in this chapter, namely the difference between *measuring* a flow and *inferring* one. A blower door measures a flow directly, being the fan flow required to sustain a pressure difference, and then applies a leakage model to characterise the envelope. A tracer experiment measures a concentration and infers transport through the mass balance of Equation 2.1. A sensor network measures concentrations at discrete points and requires an airflow model to convert them into anything else. Each step away from direct measurement introduces a model, and each model introduces uncertainty that is not visible in the instrument's specification.

8.2 Application Domains

The literature concentrates on four settings, and the requirements differ enough that a method suitable for one may be impractical in another.

Table 8.2: Application domains and their measurement constraints

Domain	**Primary question**	**Dominant constraint**	**Usual approach**
Dwellings	Infiltration rate; exposure	Occupancy prevents deliberate release; long averaging needed	Passive PFT or occupant CO_2 [10], [106]
Schools	Adequacy of ventilation per occupant	High and variable occupancy; low budget	Low-cost CO_2 networks [40], [94]
Offices	Distribution and system performance	Mechanical systems; zoning	CO_2 with system data; duct tracer testing [102], [44]
Tall and complex buildings	Interzonal transport; envelope by zone	Shafts and stairs couple the zones	Zone-resolved pressurisation; multizone tracer [107]

The school case is the one with enough field measurement behind it to state what is actually found rather than what is required. Batterman and colleagues measured ventilation rates in 37 recently constructed or renovated mechanically ventilated schools in the United States, including buildings certified under LEED and Energy Star [108]. Air change rates during the school day averaged $2.0 \pm 1.3\,\mathrm{h}^{-1}$, and only 22 per cent of classrooms met the recommended minimum ventilation rate, while median CO_2 exceeded 1000 ppm in 28 per cent of classrooms and 2000 ppm in 4 per cent, with 15-minute peaks above 2000 ppm in 19 per cent. The result establishes both the scale of the measurement problem and the reason it matters, since these were recent buildings with mechanical systems and most of them were under-ventilated during use.

The corresponding control question is what a measurement is for. Merema and colleagues reviewed demand-controlled ventilation in operating school and office buildings and reported reductions of 25 to 55 per cent in fan energy and 25 to 32 per cent in ventilation heat losses relative to a constant-volume system at design flow [109]. Savings of that magnitude are the reason CO_2 sensing is installed at all, and they also set the accuracy the sensing must achieve, since a control loop that mis-estimates ventilation demand trades the saving against the exceedances of the previous paragraph.

A pattern visible across all four domains, and consistent with the corpus analysis of Section 4.2, is that system design is reported in more detail than system validation. Deployment architecture, node count and communication topology are described precisely; the accuracy of the resulting ventilation estimate frequently is not.

8.3 Envelope Leakage and Natural Infiltration

This is the most frequently compared pair of methods and the most frequently misreported. Section 2.5 established that fan pressurisation and tracer testing measure different physical states; this section examines what happens when they are compared directly.

8.3.1 Extrapolation from 50 Pa to operating conditions

A blower-door result is obtained at an imposed pressure difference, conventionally 50 Pa. Converting it to an infiltration rate at the 1–4 Pa differences that occur naturally requires an empirical relation. The simplest is the divide-by-twenty rule, $\lambda \approx ACH_{50}/20$; more developed forms such as the Sherman correlation introduce coefficients for climate, building height, wind shielding and envelope tightness.

Patel and Mitsingas applied both to a highly airtight residential building and compared them against CO_2 tracer decay [100]. The measured envelope leakage was $ACH_{50} = 0.590\,\mathrm{h}^{-1}$, satisfying the Passive House criterion of $0.6\,\mathrm{h}^{-1}$. The divide-by-twenty rule gave an estimated $0.030\,\mathrm{h}^{-1}$ at operating conditions; the Sherman correlation, with coefficients for a 1.5-storey building in poor wind shielding with a tight envelope, extrapolated to 4 Pa and gave $0.028\,\mathrm{h}^{-1}$.

The tracer results are where the comparison becomes informative. Decay testing with the mechanical system operating gave a weighted average of $0.082\,\mathrm{h}^{-1}$ across three rooms, but that figure combines natural infiltration with mechanical ventilation. Separating the two using the alternating slopes of the decay curve gave $0.171\,\mathrm{h}^{-1}$ attributable to full ventilation mode and **$0.032\,\mathrm{h}^{-1}$ attributable to natural infiltration** [100].

The agreement between the two approaches is therefore close, with 0.028 and 0.030 from extrapolation against 0.032 from tracer decay, but only once the mechanical contribution has been separated from the tracer result. Compared without that separation, the tracer figure of $0.082\,\mathrm{h}^{-1}$ exceeds the extrapolated values by a factor of nearly three, and a reader given only those two numbers would conclude that the methods disagree badly. They do not, because they were measuring different things.

8.3.2 Low-pressure alternatives

Vega Pasos and colleagues compared tracer methods against both a conventional blower-door test at 50 Pa and the pulse technique, which measures air leakage near 4 Pa and therefore much closer to the operating regime [101]. Testing a detached two-storey house, they applied CO_2 in both constant-concentration and decay

configurations alongside the two pressurisation methods, with each test repeated under varying weather.

The result of interest is that the pulse technique reduces the extrapolation problem rather than eliminating it. Measuring nearer the operating pressure shortens the extrapolation and lowers the associated uncertainty, but the pulse method proved itself sensitive to high wind, which is the same environmental variability that makes tracer results time-specific. Moving the measurement closer to operating conditions moves it closer to the conditions that make operating measurements variable.

8.3.3 Defensible comparison requirements

The evidence supports four requirements. Tests should be repeated rather than performed once, because the natural-infiltration quantity itself varies, and repeated decay tests in dwellings have shown coefficients of variation from 20 to 64 per cent [14]. The building configuration must be identical across methods, including window position, mechanical operation and internal doors. Weather must be documented, since both the tracer result and the low-pressure pressurisation result depend on it. And the comparison must be made on a common output, being an operating-condition air-change rate, rather than by placing ACH_{50} beside a tracer λ without stating the pressure difference and the model connecting them.

8.4 Tall Buildings and Interzonal Leakage

Tall buildings introduce a problem the single-zone framework of Chapter 2 cannot represent, because stairs, lift shafts and service risers couple the floors, so leakage measured at one zone includes flow from adjacent zones as well as from outdoors.

Bahnfleth and colleagues examined this directly in developing a field protocol for tall-building envelope testing [107]. Floor-by-floor pressurisation proves difficult because adjacent zones must themselves be controlled or pressurised; otherwise interzonal flow through the shafts is measured as envelope leakage and the result overstates the fabric's permeability. Their proposal was to use the building's existing air-handling equipment to pressurise a zone, which avoids the logistical problem of deploying blower doors across many floors.

The tracer route to the same problem is to use several tracers at once, releasing a distinct gas in each zone so that the resulting concentration matrix identifies where air in any given zone came from. Fisk and colleagues built such a system around six gases, sulfur hexafluoride and five halocarbons, analysed by gas chromatography with electron capture detection, with programmable stand-alone injectors and samplers so that many locations could be sampled without an instrument at each [51]. Its accuracy was established in two ways that are worth separating. Four of the

six tracers returned the same air exchange rate, age of air and ventilation efficiency within about 15 per cent of one another, which is an internal-consistency result, and in a well-mixed test space those same four agreed within 12 per cent of a physically measured reference, an orifice-plate flow meter, which is a reference-agreement result. The two other tracers did not track the group, so a multi-tracer method also has to establish which of its tracers are behaving.

Applied to occupied commercial buildings, the same approach yielded the distribution the single-zone model cannot see [110]. Within regions served by a single air handler, local ages of air varied by 30 per cent or less from the region average, different tracers returned the same local age within 10 per cent, and duplicate samplers at one location agreed within about 5 per cent. In a well-mixed chamber with a physically measured ventilation rate, the maximum deviation of the tracer result was 12 per cent. The study also identifies a structural limitation that no improvement in analysis addresses, which is that the infiltration fraction is obtained as the difference between unity and a sum of measured fractions that is itself close to unity, so its uncertainty is large whenever infiltration is small. That is the same cancellation problem, in a different guise, that Section 9.2 identifies in multizone inversion.

The high-rise case has been examined with both approaches simultaneously. Reardon and colleagues measured air change rates with SF_6 while tracking occupant-generated CO_2 throughout a large high-rise office, repeating the tests at outdoor-air damper settings of 100, 75, 50, 35 and 20 per cent [111]. Three findings transfer. Mixing was complete within 60 to 80 minutes after an unbalanced injection and, contrary to expectation, its speed did not depend on the ratio of recirculated to outdoor air, so a mixing time established at one damper setting held across the range. CO_2 measured at the top of the main return shafts represented the whole-building average to within a correction of 2 to 4 per cent, which identifies the return plenum as a defensible single sampling point for a whole-building measurement. And the daily peak and average CO_2 concentrations correlated well with the inverse of the air change rate, as the steady-state relation predicts. Set against the placement sensitivity of Section 7.9, the contrast is instructive, since within an occupied room position dominates, whereas at the return of a well-mixed system one point suffices.

The general implication extends beyond tall buildings. The more internally connected a building is, the less a single whole-building leakage figure means diagnostically. Any pressure or tracer test should state the boundary of the zone being measured and whether leakage through shafts, doors and duct penetrations is included in the reported figure. Where that boundary is not stated, results from different buildings are not comparable even when the measurement procedure was identical.

8.5 Ductwork and Air-Distribution Leakage

Tracer methods apply on the mechanical side of the building as well as to the envelope, and the measurement problem is better posed there because the flow is confined and fan-driven rather than natural and variable.

A known tracer dose is introduced into a supply, return or outdoor-air stream and the concentration measured downstream. For a configuration handling 100 per cent outdoor air, the volumetric flow follows from the injection rate and the downstream concentration,

$$Q = \frac{G}{C_{down} - C_{up}} \tag{8.1}$$

where G is the tracer injection rate and C_{down} and C_{up} are the concentrations downstream and upstream of the injection point. Recirculating systems require the recirculated fraction to be accounted for, since tracer returning through the recirculation path raises C_{up} and biases the estimate downward if ignored.

Cheong examined tracer dilution as an alternative to conventional velocity-traverse methods for balancing air-distribution systems, where vane anemometers and Pitot tubes require knowledge of duct cross-section and are difficult to apply in complex ductwork [102]. The tracer approach measures the flow directly and is insensitive to duct geometry, which is its principal advantage.

The binding requirement is uniform cross-sectional mixing. Equation 8.1 assumes C_{down} represents the whole duct flow rather than a filament of concentrated tracer that has not yet dispersed, and injection and sampling arrangements must therefore provide sufficient residence length or deliberate mixing elements between them. A single downstream sensor can estimate total flow when this condition is met and will produce a systematically wrong answer when it is not, and the error is not detectable from the measurement itself, which is why the mixing length must be designed rather than assumed.

A related constraint applies to diagnosing leakage by difference. Where the objective is to identify loss between two points, the measurement uncertainty at each point must be smaller than the flow difference being sought. A leakage of ten per cent cannot be diagnosed with an instrument whose flow estimate carries fifteen per cent uncertainty, however carefully the tracer is dosed.

8.6 Contaminant-Source Localisation

Source localisation inverts the usual problem. Rather than releasing a known tracer to estimate airflow, the network observes concentration transients and infers the location, strength and timing of an unknown release. It is the application in

which sensor networks contribute most and in which their limitations are most consequential.

8.6.1 The forward model and why inversion is hard

The concentration at sensor i arising from a set of sources j can be written

$$C_i(t) = \sum_j S_j \, h_{ij}\Big(t - \tau_j\Big) + \varepsilon_i(t) \tag{8.2}$$

where S_j is the strength of source j, h_{ij} is the transport response from source j to sensor i, τ_j is the release time, and ε_i is measurement and modelling error. Localisation requires recovering S_j, τ_j and the source position from observations of C_i.

The difficulty is that h_{ij} is not a property of the building alone. It depends on airflow, mixing, door and window position, the state of the mechanical system, and sensor placement, all of which may change during the release being localised. The inverse problem is therefore ill-posed in a specific way, in that the transport response which must be known to invert the measurement is itself the quantity the rest of this thesis has shown to be difficult to determine.

8.6.2 Model-based approaches

Three approaches in the corpus address this differently, and the differences are instructive.

Sohn and colleagues used a Bayesian Monte Carlo formulation, precomputing a library of release scenarios spanning source location, strength, duration, weather and building condition, with incoming sensor data updating the posterior probability of each scenario [103]. Two findings are practically significant, namely that concurrent sampling across sensors identified the source faster than sequential sampling, and that noisier data required proportionally more observations to reach the same confidence. The method shifts computational cost to a precomputation stage, which suits real-time operation.

Cai and colleagues used transient accessibility functions generated by computational fluid dynamics, with an online linear program estimating source strengths and a separate procedure inferring the unknown release time from the first responding sensor [104]. In a simulated office with six candidate source locations and five sensors, two simultaneously active sources were recovered exactly from error-free data. The reliance on a CFD-derived response is both the method's strength and its limitation, since it is accurate where the model is accurate.

Vukovic and colleagues took a reduced-order approach, training a neural-network locator on fitted concentration-response coefficients rather than raw time series,

with a growing-and-pruning procedure selecting sensor positions [105]. In a three-zone test building the source zone was identified with complete accuracy under the reported validation case; in a larger 23-zone office model, three sensors were sufficient to reach approximately 80 per cent accuracy.

8.6.3 Sensor placement as part of the inverse problem

The Vukovic result carries the most general implication in this section. Placement selection was not a deployment convenience but part of the algorithm, and the accuracy achieved with three well-chosen sensors in a 23-zone model indicates how much position contributes relative to node count.

This reframes a decision usually treated as secondary. Sensors near supply diffusers, returns, doorways or stagnant corners respond very differently to the same release, so h_{ij} in Equation 8.2 varies by orders of magnitude across candidate positions. A sparse network at informative locations paired with a sound airflow model can outperform a dense network placed by convenience, which is consistent with the placement findings reported for demand-controlled ventilation in Section 6.2 [44].

8.6.4 Apportionment as the achievable form of the problem

The three methods above solve the inverse problem from a transport model. A separate literature attacks the same question statistically, apportioning measured concentrations to sources by their covariance structure rather than by modelling the airflow. Higgins and colleagues, reviewing low-cost sensing for source apportionment, report positive matrix factorisation as the most used receptor method at 31 per cent of studies, principal component analysis at 27 per cent and chemical mass balance at 7 per cent, and argue that receptor models suit indoor measurement better than source-oriented models precisely because they need only concentration data rather than an emissions inventory or meteorology [112]. Where the sources are chemically distinguishable this can succeed on low-cost data alone, since one reviewed deployment attributed 31 to 55 per cent of $PM_{2.5}$ to solid-fuel burning, and another attributed 95 per cent of ultrafine particles to outdoor origin.

The same review states the constraint on transferring this to ventilation work. Measurement accuracy for NDIR CO_2 is typically ± 30 to 50 ppm against a reference, and where the quantity of interest is a small concentration *difference* of comparable size, the data are not usable for it [112]. That is the boundary between the two uses of the same hardware. Ranking sources needs contrast between sensors; inferring a flow needs a difference the instrument can resolve.

Two further findings from that review bear on how any of this should be reported. Chemical mass balance requires near-complete source profiles and known

experimental uncertainties, which low-cost data frequently do not carry; and of the studies reviewed, only 20 per cent stated the building's construction or renovation year, 25 per cent the building type and 50 per cent the floor on which measurements were taken. The context needed to transfer a result is routinely absent, which is the same deficiency identified from the corpus as a whole in Section 10.3.

8.6.5 What low-cost networks can and cannot localise

Low-cost networks are well suited to detecting that a concentration has changed, ranking likely source regions, and identifying transport pathways between zones. They are not suited to claiming precise source coordinates unless the network has been calibrated against the building's airflow state at the time of the release.

A related limitation is temporal rather than spatial. Reviewing real-time sensing of indoor pollution, Wang and colleagues found sampling periods ranging from under a day to over two years with a median of seven days [113]. A week is long enough to characterise a source and far too short to characterise a building, whose ventilation varies with weather, season and occupancy, or to observe the sensor drift of Section 9.3. Localisation results drawn from campaigns of that length describe the airflow state that happened to obtain while the network was installed.

The limitations follow directly from Chapter 6, in that cross-sensitivity and drift corrupt C_i, differing response times between nodes distort the temporal information on which τ_j depends, packet loss and clock misalignment break the synchronisation that multi-sensor inversion assumes, stratified or recirculating flow violates the transport model, and every model-based method inherits the uncertainty of the precomputed airflow or CFD model it rests on. A localisation result should therefore be reported as a probability field over candidate positions rather than as a coordinate, which is what the Bayesian formulation produces natively and what the other approaches can be made to produce.

8.7 Method Selection by Diagnostic Goal

Table 8.3 states which method the evidence supports for each diagnostic goal, with the condition that must hold for the result to be defensible.

Table 8.3: Method selection by diagnostic goal

Goal	**Preferred method**	**Condition for a defensible result**
Compare envelope tightness between buildings	Fan pressurisation at 50 Pa	Report ACH_{50} as such; do not extrapolate without stating the model [100]
Infiltration under operating conditions	Tracer decay, or pulse testing near 4 Pa	Repeat under documented weather; separate mechanical from natural contribution [100], [101]
Long-term average air change	Passive PFT sampling	Source emission characterised over the deployment temperature range [10]
Ventilation adequacy in occupied rooms	Low-cost CO_2 network	Placement justified; background corrected; treated as screening not measurement [94], [44]
Duct flow and distribution leakage	Tracer dosing with downstream sampling	Mixing length designed; uncertainty smaller than the difference sought [102]
Zone-resolved leakage in tall buildings	Zone pressurisation with adjacent-zone control	Zone boundary stated; interzonal paths accounted for [107]
Source localisation	Sensor network with a model-based inverse method	Placement optimised as part of the method; result reported as a probability field [103], [104], [105]

Two observations close the chapter. The first is that low-cost networks appear in this table as a complement rather than a substitute in every row, since they extend the spatial and temporal coverage of a measurement whose absolute scale is set elsewhere, which is the same conclusion the platform architecture of Section 7.9 reaches from the hardware side.

The second is that the condition column is where most reported results fail rather than the method column. The methods are established and the physics is understood; what distinguishes a defensible diagnostic result from an indefensible one is whether the mixing length was designed, the zone boundary stated, the weather documented, the mechanical contribution separated, or the placement justified. The following chapter takes up the question of what a validation protocol adequate to these conditions would have to contain.

Chapter 9

Calibration, Validation and Uncertainty

Section 7.8 established that the platforms in the corpus use at least five different meanings of the word validation, and that only one of them tests the quantity a tracer measurement exists to produce. That was a survey question, asking what these systems have demonstrated, whereas this chapter addresses the complementary question of what a validation adequate to the claim would actually have to contain.

The distinction matters because the two are frequently conflated. A system that has been calibrated against certified gas, shown to respond plausibly to window opening, and found consistent across its nodes has satisfied three necessary conditions and demonstrated nothing about the accuracy of the air-change rate it reports. Establishing that requires propagating error through the inference, not only characterising the sensor.

The chapter proceeds from the measurement chain outward. Section 9.1 propagates concentration and background error through the decay estimate and shows where the uncertainty is concentrated. Section 9.2 extends this to the multizone case, where the amplification is far larger than is generally appreciated. Sections 9.3 and 9.4 treat calibration and co-location as the two practical instruments available. Section 9.5 states the protocol the evidence supports.

9.1 Uncertainty in the Decay Estimate

The decay method infers the air-change rate from the slope of $\ln(C - C_{out})$ against time. For a two-point estimate over an interval $\Delta t = t_2 - t_1$, Equation 2.3 rearranges to

$$\lambda = \frac{1}{\Delta t}\,\ln\!\left(\frac{C_1 - C_{out}}{C_2 - C_{out}}\right) \tag{9.1}$$

Treating the concentration measurements as independent with standard deviation σ_C, and the background estimate as carrying its own error σ_b, first-order propagation gives

$$\sigma_\lambda^2 \;\approx\; \frac{\sigma_C^2}{\Delta t^2}\left[\frac{1}{(C_1 - C_{out})^2} + \frac{1}{(C_2 - C_{out})^2}\right] \;+\; \frac{\sigma_b^2}{\Delta t^2}\left[\frac{1}{C_1 - C_{out}} - \frac{1}{C_2 - C_{out}}\right]^2 \tag{9.2}$$

Three properties of Equation 9.2 govern the design of a decay measurement, and each has a direct consequence for low-cost instrumentation.

The uncertainty is dominated by the tail of the decay. Both bracketed terms contain $1/(C_2 - C_{out})$, and $C_2 \to C_{out}$ by construction as the tracer disappears. Concentration error at the end of the fitting window is weighted far more heavily than the same error at the beginning. Extending the fit to capture more of the decay therefore improves the estimate only until the excess concentration approaches the noise floor, after which it degrades it.

Figure 9.1 shows the resulting behaviour. The relative uncertainty falls as the fitting window lengthens, reaches a minimum, and rises again as the tail contributes more noise than signal. The optimum is an interior point, and its position depends on σ_C, since a noisier instrument reaches its minimum earlier and at a higher floor.

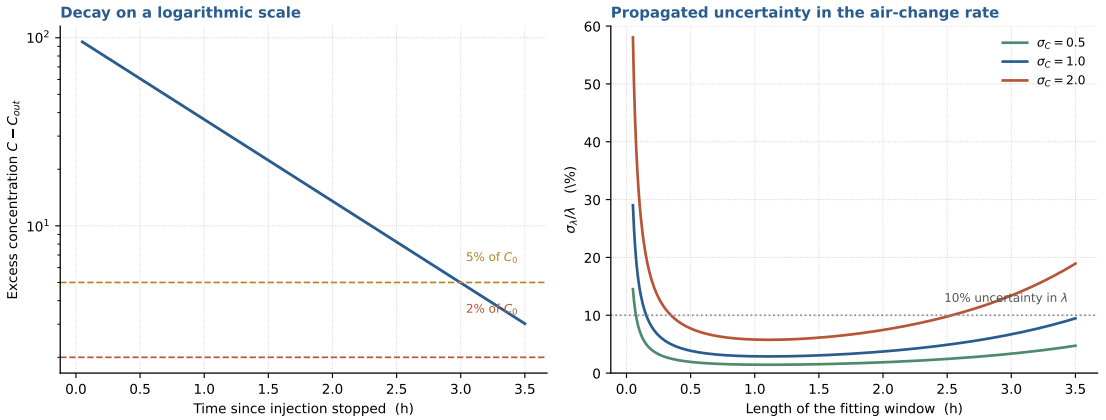


Figure 9.1: On the left, an exponential decay on a logarithmic scale, with two candidate cut-off thresholds at 5 and 2 per cent of the initial excess concentration. On the right, the relative uncertainty in λ computed from Equation 9.2 for three noise levels. Extending the fitting window reduces uncertainty until the excess concentration approaches the noise floor, after which the $1/(C - C_{out})$ term dominates and the uncertainty rises again, so the optimum is an interior point rather than the longest possible fit, and a noisier instrument reaches that optimum earlier and at a higher floor.

Background error enters differently from concentration error. The second term of Equation 9.2 is a difference rather than a sum, so an error in C_{out} that is constant across the measurement partially cancels. It does not cancel completely, and the residual grows as the decay proceeds for the same reason as above. This is why occupant-generated CO_2 measurement, where the background is both large and variable, carries an uncertainty that no improvement in sensor resolution addresses [40], [2].

The requirement is a calibration valid at the low end, not a good calibration on average. A sensor calibrated at mid-range and used across a decay is least accurate where Equation 9.2 weights it most. This is the analytical reason behind the empirical result reported for the TEBE-SENSE platform in Section 7.7, where the 2000 ppm module produced substantially greater dispersion than the 500 ppm modules and overestimated flow rates at low concentration [54]. The full-scale range must be matched to the concentration the decay will end at, not the concentration it begins at.

9.1.1 Fitting the decay

Equation 9.2 describes a two-point estimate, and the log-linear regression normally used in its place inherits the same weighting problem, because ordinary least

squares on $\ln(C - C_{out})$ assumes constant variance in the transformed variable, which Equation 9.2 says is precisely what does not hold.

Miller and colleagues encountered this directly. Rather than fitting a slope, they fitted the full concentration profiles predicted by the two-zone mass balance to measured data by nonlinear least-squares minimisation, refining the airflow estimates iteratively until model and data agreed [114]. Validated against blower-imposed flows in two rooms of a test house, the recovered airflows agreed with the imposed values within an average absolute error of 8 per cent. The methodological finding is in the residuals, since an initial fit assuming a single standard deviation across all measurements produced residuals whose spread was not constant, so the minimisation was reweighted with the measurement standard deviation modelled as a power function of the fitted value. Parameter uncertainties were then obtained by Monte Carlo simulation rather than from the covariance matrix.

That sequence is the empirical counterpart of Equation 9.2. The analysis predicts that measurement error should carry unequal weight along the decay; the fitting exercise found exactly that pattern in the residuals and had to correct for it to obtain defensible parameter uncertainties. Two practical consequences follow for a low-cost platform. An unweighted log-linear fit over the whole decay is not merely suboptimal but systematically misstates its own confidence interval, and the fitting window and weighting scheme are part of the measurement method and belong in the report alongside the sensor specification.

9.2 Error Amplification in Multizone Inversion

The single-zone case is the favourable one. Where a building is treated as several coupled zones, the airflows are recovered by inverting a system relating measured concentrations to interzonal flows, and the conditioning of that system determines how input error propagates.

D'Ottavio and colleagues performed the error analysis for perfluorocarbon-derived multizone ventilation measurements and reported the result quantitatively [115]. With source rate errors fixed at 10 per cent and average concentration errors ranging from 2 to 30 per cent, the resulting flow errors spanned **10 to 651 per cent** across systems whose condition numbers ranged from 1.00 to 58.01. The upper end of that range is not a typographical curiosity, since a badly conditioned multizone system converts a 30 per cent concentration error into a flow estimate wrong by a factor of six.

The condition number is therefore the quantity that determines whether a multizone measurement is meaningful, and it is a property of the zone configuration and the sampling scheme rather than of the instrument. The same analysis found that first-order error propagation ceased to agree with the empirical result once

concentration errors approached 50 per cent, because both schemes assume the concentration errors are small relative to the concentrations themselves, and that assumption fails precisely in the low-concentration regime identified in Section 9.1.

Sherman's analysis of passive multizone techniques reaches a compatible conclusion from the sampling side [116]. For a ranch-style house, omitting either injection or sampling from one zone introduces an additional 10 to 25 per cent uncertainty in the estimated tracer concentration and hence in ventilation. For a colonial-style house the corresponding figure is 40 to 80 per cent, the difference arising from the greater interzonal coupling of the multi-storey plan.

Two conclusions follow for low-cost platforms. First, adding nodes is not neutral, because a network that samples some zones and not others may produce a worse-conditioned system than a smaller, symmetric deployment. Second, a validation performed in a single well-mixed chamber establishes nothing about multizone performance, because the error amplification that dominates the multizone case is absent by construction from the single-zone one.

9.3 What Calibration Must Cover

Calibration in this literature is frequently reported as a single operation performed once. The evidence supports treating it as a specification with several dimensions, each of which has been shown to matter independently.

Range. Section 9.1 establishes that the low end of the range carries the most weight. A single-point or mid-range calibration leaves the influential region unconstrained.

Absolute scale. On-site evaluation of twenty-three low-cost NDIR CO_2 sensors found near-perfect linearity against a reference ($R^2 = 0.999$) but a mean root-mean-square error of 18.3 ppm before correction, with 0.45 per cent of measurements outside the manufacturer's stated accuracy [76]. Linearity and accuracy are distinct properties, and a factory calibration typically delivers the first without the second.

Range, again, and in practice. The same study found the corrected error to be range-dependent, at 7.4 ppm for ambient concentrations and 11.9 ppm below 1500 ppm, with performance declining above 1500 ppm and during periods of changing occupancy [76]. A review of calibration strategy for low-cost particulate sensors reaches the same conclusion from the opposite end of the range, reporting that linearity degrades and absolute error rises at elevated concentrations, and arguing that performance must be reported by range and the calibration model made range-specific rather than global [82]. Both results say that a single error figure quoted for a sensor is a statement about the range it was tested over, and that the range which matters for a decay is set by Section 9.1 rather than by the manufacturer.

Environmental conditions. Response depends on temperature and humidity for every principle reviewed in Chapter 6. A calibration performed at laboratory conditions does not transfer to a deployment that spans a heating season.

Time. Metal-oxide sensing layers change irreversibly during operation, so a model trained on an initial calibration degrades as the sensor ages; Baur and colleagues evaluated calibration strategies against this drift directly and found periodic recalibration or explicit drift compensation necessary to maintain usable output [70]. Electrochemical cells consume their electrolyte over a service life measured in months to a few years. A deployment intended to run for a year requires a recalibration schedule, not an initial calibration.

Unit-to-unit variability. Specifying a sensor model does not specify its performance. Ahn and colleagues evaluated 264 low-cost dust sensors against a controlled decaying particle concentration and found them to separate into four distinct accuracy classes [69]. A deployment that assumes interchangeability between nominally identical units is assuming something the evidence contradicts, and per-unit calibration or cross-calibration against a common reference is required.

9.4 Co-location as the Validation Instrument

Co-location, meaning the operation of the device under test alongside a reference instrument in the same air, is the practical mechanism by which the top rung of the validation ladder in Figure 7.1 is reached. Recent work has established what such a study must report.

9.4.1 Separating precision from accuracy

Zaky and colleagues evaluated two consumer-grade monitors for CO_2 and $PM_{2.5}$ against two reference instruments, using time-series analysis, linear regression, Pearson correlation, root-mean-square error, Bland–Altman limits of agreement and paired t-tests [90]. The finding that matters methodologically is that one device showed a higher coefficient of determination, stronger correlation and narrower limits of agreement, yet a larger mean difference and RMSE than the other. It tracked relative change better while reporting absolute values worse.

No single accuracy figure can express this, because the two properties are separately relevant. For demand-controlled ventilation, tracking relative change is what the control loop requires, whereas for a decay fit the ratio $(C_1 - C_{out})/(C_2 - C_{out})$ in Equation 9.1 depends on relative accuracy, and a background-corrected absolute concentration is needed to establish C_{out} itself.

9.4.2 Co-location benefits, and what it does not require

Co-location is usually presented as a substitute for a traceable calibration that the study could not afford. One result in the corpus reframes it as the appropriate instrument in its own right. Honan and colleagues co-located twenty-three low-cost NDIR modules over three test periods in Irish school conditions and derived correction factors from the co-location itself, without laboratory-grade reference equipment [76]. Correction reduced the root-mean-square error by 27 per cent, from 18.3 to 13.4 ppm, mean proportional bias by 31 per cent, from 68 to 46 ppm, and out-of-range measurements by 43 per cent.

The argument behind the method is the part worth carrying forward. A ventilation rate inferred from CO_2 depends on the indoor–outdoor concentration *difference*, not on either absolute concentration, so what the network must achieve is mutual consistency on a common scale rather than individual traceability. Where that is the quantity of interest, a network-internal reference is a defensible substitute for certified calibration. The qualification is that this reasoning does not extend to a deliberate-tracer measurement, where the background is near zero and the absolute concentration is the measurand; there the argument of Section 9.1 applies and traceability cannot be dispensed with.

9.4.3 The size of the correction

How much co-location changes depends on the pollutant, and the magnitudes are large enough to determine whether a quantitative claim is possible at all. Tryner and colleagues built nine identical multi-sensor enclosures and co-located them with reference monitors in the kitchen of an occupied house for 168 hours [117]. Interpreted with the manufacturers' calibrations, the units classified each pollutant into the same broad band as the reference and identified the same dominant sources, but ozone showed no correlation with the reference at all. Re-interpreted with linear mixed calibration models fitted from the co-location, the root-mean-square error for CO_2 fell from 408 to 58 ppm, for nitrogen dioxide from 22 to 14 ppb and for ozone from 21 to 7 ppb, and the ozone correlation rose from $r = 0.02$ to $r = 0.75$.

Two details of that model are instructive. It required a random intercept and slope per sensor serial number, because the individual units did not share a response, which is a direct confirmation of the per-unit requirement in Section 9.3 arrived at from the modelling side. And a sevenfold reduction in CO_2 error establishes the order of magnitude at stake, since a factory-calibrated module carrying a 408 ppm error cannot support a decay fit under Equation 9.2, whereas the same hardware after co-location can. The distinction between a screening device and a measuring instrument is, in this instance, entirely a matter of calibration procedure.

9.4.4 Laboratory agreement does not predict field agreement

The strongest reason to require co-location in the deployment environment is that agreement established in a chamber has repeatedly failed to survive the move indoors. Moreno-Rangel and colleagues, evaluating a low-cost monitor for residential exposure assessment, assembled the comparison across devices, and one monitor reached $R^2 = 0.90$ against a reference in controlled chamber experiments and $R^2 = 0.86$ in residential settings, while another that performed well in the laboratory fell to $R^2 = 0.3$ indoors and $R^2 = 0.1$ to 0.2 outdoors at low concentrations, overestimating $PM_{2.5}$ by 200 per cent indoors and 500 per cent outdoors [118]. The device they evaluated themselves had been characterised only in the laboratory, where it correlated at $r = 0.99$.

The pattern is consistent with the range argument of Section 9.3, since chamber tests are conducted at concentrations convenient to generate, and the failures appear at the low concentrations that dominate real indoor air and, by Equation 9.2, the tail of a decay. A laboratory correlation coefficient is therefore evidence about the laboratory.

9.4.5 Reliability as a measurable quantity

A co-location study of 87 low-cost environmental monitors found that 5 of the 87 units failed outright during the exercise [89]. A failure rate near six per cent is a design parameter for a dense network rather than an anomaly, since a deployment of twenty nodes should expect to lose one, and an analysis that requires all nodes to report will fail. Reliability therefore belongs in the validation report alongside accuracy.

9.4.6 Laboratory protocols

Structured laboratory evaluation has matured faster than its use. Protocols now exist for assessing low-cost sensors under controlled laboratory conditions and under conditions simulating real environments [91], [92], and for evaluating large populations of units against a controlled decaying concentration [69]. Controlled-source testing with known emissions from cooking, cleaning, candles and cigarettes has been used to establish whether devices detect the events they would need to detect in service [119].

9.5 A Validation Protocol

Table 9.1 sets out what a validation sufficient to support a ventilation claim would contain, with the reason each element is necessary and the consequence of omitting it. It is constructed from the failures identified in this chapter and in Section 7.8 rather than from an idealised standard.

Table 9.1: Elements of a validation adequate to a ventilation claim

Element	Requirement	Consequence of omission
Range calibration	Cover the full range the decay traverses, including the tail	Largest error where Equation 9.2 weights it most [54]
Environmental span	Calibrate across the deployment temperature and humidity range	Systematic seasonal bias, undetectable from the data [70]
Per-unit characterisation	Calibrate or cross-calibrate each unit	Inter-unit spread mistaken for spatial variation [69], [89]
Recalibration schedule	Defined interval with drift compensation	Slow degradation reported as a change in the building [70]
Reference co-location	Simultaneous reference measurement of the same event	No basis for an absolute claim; only the lower rungs of Figure 7.1 [90]
Separated error reporting	Concentration error, decay-fit uncertainty, spatial variability and reference agreement reported individually	A single accuracy figure conceals which term is limiting [90]
Repetition	Repeated tests under documented conditions	The measurand itself varies; coefficients of variation of 20–64% have been reported [14]
Zone and conditioning statement	Zone boundary defined; multizone conditioning assessed	Error amplification of up to two orders of magnitude goes unreported [115], [116]
Reliability reporting	Node failure and data-loss rates over the campaign	Analyses assume complete data that a real network does not deliver [89]

The table specifies what must be reported. A parallel body of work specifies when each check should occur, and the two are complementary. Feng and colleagues set out a staged quality-assurance workflow for low-cost particulate networks running from procurement screening, through bench checks and co-location with temporally blocked validation, to external validation on data the calibration never saw, and

finally to rotating "gold-node" drift checks in which units are cycled back to the reference during the campaign [82]. Two features distinguish it from a calibration schedule. Blocked and external validation test whether a correction model transfers rather than whether it fits, which is the failure mode behind the laboratory-to-field collapse of Section 9.4; and rotating the reference through the network measures drift during deployment rather than inferring it afterwards from opening and closing co-locations.

The same review identifies the reporting gap that Table 9.1 is intended to close, which is that uncertainty is seldom carried forward from the calibration to the quantity a decision actually rests on, and that the boundaries of applicability, covering which particle types, which concentration ranges and which environmental conditions, are rarely stated explicitly [82]. Transposed to ventilation measurement, the quantity a decision rests on is the air-change rate, and carrying uncertainty forward to it is exactly the propagation of Section 9.1 that the corpus does not report.

Two observations close the chapter. The first is that no platform in the corpus satisfies this table. The strongest, examined in Chapter 7, satisfy the calibration and reference-co-location rows under laboratory conditions and are silent on the remainder. This is stated as a description of the field rather than a criticism of individual studies, since the elements are drawn from nine separate sources and no single study had reason to address all of them.

The second is that the elements are unequally expensive. Range calibration, separated error reporting, repetition and a zone statement cost analysis time rather than equipment, and their absence is the most common and least excusable gap. Reference co-location over a full deployment period is genuinely costly, and it is the element whose absence most limits what the field can currently claim. The experimental programme proposed in the final chapter is constructed around supplying it.

Chapter 10

Synthesis and Conclusions

This chapter draws the review together. Section 10.1 answers each research question from the evidence assembled in Chapters 5 to 9. Section 10.2 states the strengths the literature establishes and Section 10.3 the limitations it exposes. Section 10.4 sets out the measurement architecture the evidence supports, and Section 10.5 states the research gap in the form of the specific study that would close it.

10.1 Answers to the Research Questions

10.1.1 RQ1. Landscape and maturity

Low-cost indoor sensing is not a single technology at a uniform stage of development, and the phrase should not be used without naming a target. Maturity divides cleanly by detection principle. Non-dispersive infrared sensing of carbon dioxide and optical scattering measurement of particulate matter support quantitative use after appropriate calibration; a controlled assessment found carbon dioxide agreement within 15 per cent while the same devices produced poor quantitative agreement for other pollutants measured simultaneously [67]. Metal-oxide volatile organic sensing supports event detection and relative comparison but reports a proxy rather than a concentration, because the sensing layer responds to whatever adsorbs onto it [65], [70].

Tracer-gas methodology is mature in a different sense, being settled rather than developing. The mass balance, the three test designs and the associated measurement standards were established between roughly 1980 and 1995 and have required no substantial revision. The corpus reflects this directly, with 71 of 226 records falling before the primary date window and retained under the seminal-methodology exception.

The maturity of the intersection is the significant result. Within the 226-record

analysis corpus, 98 records address tracer-gas methods and 74 address low-cost sensing, but only nine address both, and of those nine three name the detection principle used while two report a comparison against a reference instrument. The sensor-technologies branch of the keyword taxonomy is the thinnest in the corpus at 34 records, because most of the literature reports what was measured without specifying how.

10.1.2 RQ2. Difficulties

The technical difficulties are consistent across every sensing principle examined, and the pattern is that the transducer is rarely the limiting element. The reported limitation is drift, cross-sensitivity, environmental dependence, unit-to-unit variability, or the absence of a calibration valid under deployment conditions. Ahn and colleagues found 264 nominally comparable low-cost dust sensors to separate into four distinct accuracy classes, so that specifying a sensor model does not specify its performance [69].

The more consequential difficulties are not instrumental at all. Section 9.1 shows that uncertainty in a decay-derived air-change rate is dominated by the tail of the decay, where the excess concentration approaches the noise floor, and Section 9.2 shows that multizone inversion turns concentration errors of 2 to 30 per cent into flow errors spanning 10 to 651 per cent, with the position within that range set by the conditioning of the zone system rather than by the quality of the concentration measurement [115]. Gao and colleagues quantified a third, finding that ventilation estimates from different tracers can differ by a factor of two to four when dispersion is not passive [1]. None of these is addressed by a better detector.

The non-technical difficulties are dominated by the environmental burden of the reference tracers. Sulfur hexafluoride achieves its detectability through the chemical stability that also gives it a very high global warming potential, and Section 5.2 argues that this correlation is structural rather than accidental, because the property that makes a good tracer is the property that makes it persist.

10.1.3 RQ3. Trends

Three directions are visible in the corpus and are unevenly supported by evidence.

Networked deployment is the best established. Architecture, node count and communication topology are reported in detail across the low-cost literature, and the platforms in Section 7.3 demonstrate that dense deployment is technically routine.

Machine-learning calibration appears in 11 records and is largely applied to drift compensation rather than to the measurement itself. Baur and colleagues found

that a model trained on an initial calibration degrades as a metal-oxide sensor ages, and evaluated compensation strategies against that degradation [70].

Emerging low-cost tracer methods are the least developed and the most relevant to this thesis. Three routes are demonstrated, namely non-dispersive infrared detection of sulfur hexafluoride [54], photoionisation detection of a deliberate volatile organic tracer [49], and semiconductor detection of a refrigerant tracer for dense spatial mapping [50]. A fourth route, using passive tracers co-analysed with the volatile organic samplers a study already deploys, attacks the cost structure rather than the hardware and is the most promising unproven direction [58], [18].

10.1.4 RQ4. In-situ challenges

The distinction between laboratory and occupied-building evidence governs the entire assessment of Chapter 7, and it does not follow cost or sophistication. The two most technically developed platforms in the corpus are laboratory demonstrations; the systems with genuine occupied-building records are the simpler carbon dioxide devices.

Four challenges are specific to real buildings. Mixing is never ideal, and the single-zone idealisation on which Equation 2.3 rests fails progressively as a building becomes more compartmentalised, with Sherman quantifying the penalty for incomplete zone coverage as 10 to 25 per cent additional uncertainty in a ranch-style house and 40 to 80 per cent in a colonial-style plan [116]. Airflow varies during the measurement, so that repeated decay tests in dwellings have shown coefficients of variation from 20 to 64 per cent [14]. Background is neither small nor stable for the tracer most used in occupied buildings. And sensor placement changes the answer, since position has been shown to alter the outcome of carbon-dioxide-based ventilation control directly [44].

10.1.5 RQ5. Tracer selection

No tracer is universally optimal, and Figure 5.1 shows why, in that detectability and burden of use are correlated in the wrong direction and no candidate occupies the favourable corner. Selection is therefore a decision to be justified rather than a preference.

The evidence supports sulfur hexafluoride for short, controlled tests requiring sensitivity or comparability with historical data; perfluorocarbon tracers for multi-day multizone field averages; and carbon dioxide for occupied buildings where low cost and dense deployment matter more than a controlled source term. Among the alternatives to these, isobutene with photoionisation detection is the only one with an occupied-building reference comparison, agreeing with a simultaneous sulfur hexafluoride measurement near $1\,\mathrm{h}^{-1}$ at roughly one-fifth of the cost [49]. Stable

krypton has been validated as a replacement in standardised containment testing and avoids the greenhouse-gas objection entirely, but requires specialised detection [31]. Refrigerant tracers extend spatial coverage rather than improving accuracy [50].

10.1.6 RQ6. Applications and domains

Chapter 8 establishes that what is usually grouped as ventilation measurement comprises four distinct physical problems, measured at different pressure regimes against different reference methods. Requirements differ correspondingly by domain, since dwellings favour passive or occupant-generated methods because occupancy precludes deliberate release, schools favour low-cost carbon dioxide networks because occupancy is the variable of interest, offices require distribution-system testing alongside room measurement, and tall buildings require zone-resolved methods because shafts couple the floors [107].

The recurring failure in this material is not method selection but reporting. The comparison of Patel and Mitsingas illustrates it precisely, since extrapolated blower-door values of 0.028 and 0.030 h^{-1} agree closely with a tracer-derived natural infiltration rate of 0.032 h^{-1}, but only once the mechanical contribution has been separated from the tracer result, and compared without that separation the same data appear to disagree by a factor of nearly three [100].

10.2 Strengths the Literature Establishes

Four capabilities are demonstrated well enough to be relied upon.

Distributed measurement at a density no reference instrument permits is now routine, and it resolves spatial structure that single-point measurement averages away. The refrigerant grid of Section 7.4 sampled at ten-second intervals across multiple locations simultaneously [50].

Continuous long-duration monitoring in occupied buildings is established for carbon dioxide, with school and classroom deployments running over extended periods and producing operationally actionable findings about ventilation practice [94], [93].

The analytical feasibility of low-cost detection has been demonstrated for each of the three tracer pathways identified under RQ3. These are real results and should not be understated, since inexpensive infrared detection of sulfur hexafluoride to within ±20 per cent for rates between 1 and 7 h^{-1} is a substantive engineering achievement [54].

The methods for validating low-cost sensors have matured faster than their use. Co-location protocols now separate precision from accuracy using regression, correlation, root-mean-square error, Bland–Altman limits and paired testing [90],

and laboratory procedures exist for evaluating large populations of units against controlled concentrations [69], [91].

10.3 Limitations and Research Gaps

Context dependence is the principal limitation. Every quantitative result in this review is conditional on the building, the mixing state, the weather and the sensor position at the time of measurement. A figure reported without those conditions is not transferable, and much of the literature reports without them. The deficiency has been measured directly rather than merely observed, since in a review of low-cost sensing for source apportionment only 20 per cent of the articles stated the building's construction or renovation year, 25 per cent the building type and 50 per cent the floor on which measurements were made [112]. Half the literature does not record which storey of the building it describes.

An independent review reaches a compatible conclusion. A systematic review of low-cost sensors for indoor air quality published in 2026, using a different search strategy and different inclusion rules, screened 2736 records and retained 24 that met its accuracy and quality criteria, and of those 24 only four satisfied the technical and legal requirements for institutional or regulatory use [68]. The absolute numbers are not comparable with those reported in Chapter 4, since the two reviews scope the field differently and that review does not address tracer methods at all. The shape is comparable, and it is the same shape, being a large retrieved literature, a small fraction surviving a quality filter, and a smaller fraction again meeting the standard required for the measurement to be relied upon. That two reviews conducted independently arrive at the same structure supports reading it as a property of the field rather than an artefact of either protocol.

Validation is systematically weaker than system design. Section 7.8 identifies five distinct meanings of validation in use, of which only reference-instrument agreement tests the quantity a tracer measurement produces. Two records in the corpus report such a comparison. Section 9.5 sets out nine elements a sufficient validation would contain, and no platform in the corpus satisfies them.

The two literatures remain largely separate. The bimodal publication distribution of Section 4.1 is not a historical curiosity but a live structural feature, in that the classical tracer work and the low-cost sensing work have different authors, venues and concerns, and the trough between them, of 21 records across 1996 to 2008, marks a genuine discontinuity in the transfer of methodological knowledge.

The intersection has not grown with the field. Extending the corpus by citation chasing, the search route most likely to surface work at an intersection, added 19 records and produced no additional papers at the intersection. The nine remain nine. This indicates a property of the field rather than a limitation of the

search strategy.

Reporting practice is the most tractable gap. Of the nine validation elements in Table 9.1, five cost analysis time rather than equipment, being range calibration, separated error reporting, repetition under documented conditions, a zone and conditioning statement, and reliability reporting. Their routine absence is the least excusable feature of the current literature and the easiest to remedy.

10.4 The Measurement Architecture the Evidence Supports

The evidence converges on a three-layer architecture rather than on a replacement instrument, and each layer answers a different requirement.

Table 10.1: Three-layer measurement architecture

Layer	Instrumentation	What it contributes
Reference concentration	Photoacoustic spectroscopy, gas chromatography, calibrated infrared analyser	Absolute concentration scale and drift detection at one location, and the only basis for a quantitative concentration claim [54], [49]
Spatial	Low-cost carbon dioxide, refrigerant, isobutene or equivalent nodes	Spatial and temporal coverage; resolves structure the reference layer cannot see [50], [94]
Inference	Zone mass balances, airflow-network models, Bayesian updating, response functions, machine-learning localisation	Converts concentrations into flows, and carries its own uncertainty [103], [104], [105]

Two reference measurements sit outside the architecture rather than inside its first layer, and the distinction matters because they validate different quantities. A reference ventilation measurement, meaning an established tracer method or a direct airflow measurement, tests the output of the inference layer and is the only benchmark against which an air-change rate can be checked. Fan pressurisation is a controlled envelope benchmark measured at an imposed pressure difference, so it characterises leakage rather than the ventilation rate in the operating state, and it is not interchangeable with a tracer result, as Section 2.5 sets out and as the phased programme of Chapter 11 applies when it requires the mechanical contribution to be separated before the two are compared [100], [101].

The architecture is not novel as a proposal, since it is what the successful studies in this review implicitly did. What has not been done is validating all

three layers simultaneously in one building. Existing work demonstrates pairwise agreement, covering tracer against blower door in individual dwellings [100], [101], tracer dilution for duct and air-handler testing [102], and low-cost detection against reference analysers [50], [54], but no study closes the loop across the reference, spatial and inference layers together.

Until that evidence exists, low-cost networks are correctly treated as spatially rich complements to reference tracer methods, and to envelope testing where the envelope is the question, rather than as replacements for them. This is the position Section 8.7 reaches independently from the applications side, where low-cost networks appear as a complement in every row of the method-selection table.

10.5 The Research Gap

The gap is specific rather than general, and stating it precisely is the most useful thing this review can do.

It is not a further prototype sensor. Three detection pathways have already been demonstrated analytically, and a fourth prototype would add a fourth demonstration to a set whose limitation is not the number of demonstrations.

It is a controlled field comparison in which a low-cost tracer network operates in an occupied building alongside a reference analyser, with repeated tests across ventilation modes, concentration levels, temperatures, humidities and mixing conditions, and with concentration error, decay-fit uncertainty, spatial variability and reference agreement reported as separate quantities.

Three features of the corpus establish why this specific study is the binding one. The nine intersection records include only two reference comparisons, and both of the strongest were conducted in a laboratory layout and a 40-litre box rather than a building. The one occupied-building comparison using a deliberate tracer is nearly two decades old and used a detector that is inexpensive only relative to photoacoustic spectroscopy [49]. And the error analysis of Chapter 9 shows that the quantities such a study would need to report, being tail-weighted decay uncertainty, multizone conditioning and drift over a deployment period, are precisely those that laboratory validation cannot supply.

The experimental programme in Chapter 11 is constructed to close this gap rather than to extend the demonstration set.

Chapter 11

Future Work and Proposed Experimental Design

Section 10.5 states the research gap as a specific study, namely a controlled field comparison in which a low-cost tracer network operates in an occupied building alongside a reference analyser, with repeated tests across ventilation modes and conditions, and with error reported as separate quantities. This chapter specifies that study.

The design is prescriptive rather than descriptive, and every element in it is traceable to a finding established earlier. The three-layer architecture of Section 10.4 determines the instrumentation; the nine validation elements of Table 9.1 determine what must be reported; the error analysis of Chapter 9 determines the calibration specification and the fitting-window choice; and the four measurement problems of Section 8.1 determine the phasing, since a single experiment addressing all of them at once would confound quantities that are measured at different pressure regimes.

11.1 Objectives and Rationale

The experiment should answer four questions, stated so that each has a defined reference measurement and a pass condition.

1. Can a low-cost gas-sensor network reproduce reference measurements of operating-condition air exchange, and within what uncertainty?
2. Can tracer dilution identify flow losses or imbalance between selected duct sections, branches or air-handler paths?
3. Can the same network identify the region of a known indoor contaminant release?

4. How do mixing state, HVAC operation, weather, sensor drift and data loss affect each of the three results above?

The fourth question is the one the existing literature does not answer, and it is the reason the study must run over a period rather than as a single campaign. The two strongest reference comparisons in the corpus were conducted in a controlled laboratory layout and a 40-litre test box [50], [54], and neither setting can expose drift, material interaction, packet loss under real network conditions, or performance under naturally varying airflow.

The primary outcomes must be reported as three separate quantities, being concentration error, flow or air-change error, and source-localisation error. Reporting them together would repeat the conflation identified in Section 7.8, in which a good concentration calibration is presented as a validated airflow result.

11.2 Test Configuration and Phasing

11.2.1 Building and zone definition

The building should be divided into explicit measurement zones before any instrumentation is installed, covering the occupied or test rooms, the return-air path, the supply branches, and any adjacent spaces connected through doors, shafts or leakage paths. Section 8.4 establishes that a leakage figure without a stated zone boundary is not comparable between buildings, and Section 9.2 establishes that the conditioning of the zone system determines how input error propagates.

The zone definition should therefore be accompanied by an assessment of the conditioning of the resulting multizone system, because a network that samples some zones and not others may be worse conditioned than a smaller symmetric deployment. Where the intended analysis is single-zone, a multi-point uniformity check must precede the first decay to establish that the idealisation holds.

11.2.2 Node classes

The network implements the three-layer architecture of Section 10.4 and the platform design of Figure 7.2, with three node classes described below.

Reference nodes. One calibrated analyser, whether photoacoustic, chromatographic or a high-quality infrared instrument, in the principal test zone, and a second near the main return or supply path. These supply absolute concentration and detect network drift, and are the only basis on which a quantitative claim can be made.

Low-cost concentration nodes. Distributed sensors at room, corridor, return and supply positions. Non-dispersive infrared carbon dioxide nodes are appropriate

for occupied-building monitoring; refrigerant tracer nodes with metal-oxide detection are appropriate for exploratory spatial mapping where a deliberate release is permitted [50], [94].

Environmental and state nodes. Temperature, relative humidity, differential pressure, door state, HVAC state and outdoor weather. These are not ancillary, since Section 9.3 establishes that sensor response depends on temperature and humidity for every principle reviewed, and Section 8.3 establishes that weather must be documented for a tracer result to be interpretable.

Every low-cost node must write raw data locally as well as transmit it, clocks must be synchronised before each test, and packet loss must be logged rather than silently interpolated. The refrigerant grid of Section 7.4 demonstrates the value of ten-second sampling and, in the same study, why direct-to-cloud architecture without local buffering is fragile [50].

11.2.3 Phasing

The four measurement problems are addressed in sequence, with one network serving all phases and each phase carrying its own reference and acceptance criteria. Figure 11.1 shows the structure.

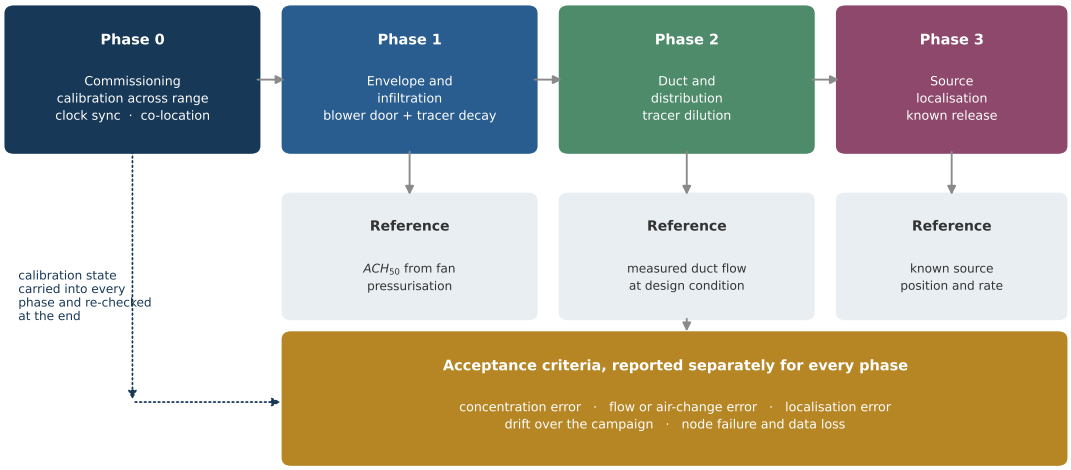


Figure 11.1: The proposed programme. One sensor network serves all four phases, commissioning establishes the calibration state carried into every phase and re-checked at the end, and each measurement phase carries its own reference measurement and its own acceptance criteria. The separated reporting row implements Table 9.1, under which a good concentration calibration must not be reported as a validated airflow result.

Phase 0, commissioning. Calibration across the full working range, co-location of every low-cost node against a reference for a defined period, clock synchronisation, and a baseline record of inter-node spread. This phase establishes the calibration state that subsequent phases assume and that the closing repeat will test against.

Phase 1, envelope and infiltration. A fan-pressurisation test with the building configuration documented and intentional openings sealed, recording the pressure–flow curve and the leakage metric at 50 Pa. This is the controlled envelope benchmark and must not be reported as the natural infiltration rate. Repeated tracer decay tests then follow under normal operating conditions, with the mechanical contribution separated from the natural one as in the comparison of Section 8.3 [100]. Repetition is required because the measurand itself varies, and coefficients of variation from 20 to 64 per cent have been reported for repeated decay tests in dwellings [14].

Phase 2, duct and distribution. Tracer dosing into a supply, return or outdoor-air stream with downstream sampling, using Equation 8.1. The injection and sampling arrangement must provide sufficient mixing length, and the flow uncertainty must be established as smaller than the leakage being sought before the test is run rather than after [102].

Phase 3, source localisation. A known release at a documented position and rate, with the network attempting recovery through the forward model of Equation 8.2. Sensor positions should be selected as part of the inverse method rather than by convenience, following the result that three well-placed sensors reached approximately 80 per cent accuracy in a 23-zone model [105]. The output should be a probability field over candidate positions rather than a single coordinate.

11.3 Calibration and Quality Control

The calibration specification follows directly from Equation 9.2 and from the nine elements of Table 9.1.

Range. Calibration must span the full range each decay will traverse, and the full-scale selection of each module must match the concentration the decay ends at rather than the concentration it begins at. This is the single most consequential specification in the chapter, and the empirical justification is the dispersion observed for the higher-range module of the TEBE-SENSE platform [54].

Fitting window. Section 9.1 shows that the relative uncertainty in λ has an interior minimum, so the fitting window should be selected from the measured noise level rather than by fitting the whole decay. The window used must be stated with the result.

Environmental span. Calibration must cover the temperature and humidity

range of the deployment. A campaign running across a heating season cannot rely on a calibration performed at a single condition.

Per-unit characterisation. Each node is calibrated or cross-calibrated individually, following the finding that 264 nominally identical dust sensors separated into four accuracy classes [69].

Recalibration and drift. A defined recalibration interval with drift compensation, and a closing co-location repeating the Phase 0 procedure. The difference between opening and closing co-location is the drift measurement, and it is the quantity no laboratory study in the corpus supplies.

Reliability. Node failure and data-loss rates recorded across the campaign. A co-location study of 87 monitors found 5 units failed outright, so an analysis requiring complete data from every node should be expected to fail [89].

11.4 Analysis and Reporting Plan

Table 11.1 states what each phase reports and the acceptance criterion proposed for it. The criteria are drawn from the performance already demonstrated in the corpus, so that the study tests whether laboratory performance survives field deployment rather than setting an arbitrary target.

Table 11.1: Reported quantities and proposed acceptance criteria

Phase	Reported quantity	Proposed acceptance criterion
0	Inter-node spread; concentration error against reference	Spread small enough to justify the intended zone model; concentration error consistent with the manufacturer specification after correction
1	Air-change rate with decay-fit uncertainty; envelope metric at 50 Pa	Agreement with the reference tracer within the ±20% demonstrated in the laboratory [54], sustained across repeated tests and weather conditions
2	Duct flow and section losses	Flow uncertainty demonstrably smaller than the leakage being diagnosed [102]
3	Probability field over candidate source regions	Correct region identified at a stated confidence, with the airflow model and its assumptions reported [103], [105]
All	Drift between opening and closing co-location; node failure and data-loss rate	Drift bounded and reported; failures reported rather than excluded [89], [70]

Three reporting rules apply across all phases.

Error quantities are reported separately and never combined into a single accuracy figure, following the finding that a device may track relative change well while reporting absolute values poorly [90]. Precision and accuracy are distinct, and a decay fit depends on the first while background correction depends on the second.

Every result carries its conditions, covering building configuration, HVAC state, weather, occupancy, sensor positions and the fitting window used. Section 10.3 identifies context dependence as the principal limitation of the existing literature, and a study that repeats the omission adds a further non-transferable result.

Negative and partial results are reported. If the network fails to reproduce the reference under some ventilation mode, that is the informative outcome, because it identifies the boundary of applicability, which is precisely what the corpus currently lacks for every platform in it.

11.5 Scope and Feasibility

The programme is deliberately bounded. It is a single-building study rather than a multi-building survey, because with the intersection literature standing at nine records a well-instrumented single case that reports its conditions completely is more useful than several cases that do not.

The instrumentation requirement is one reference analyser, a modest number of low-cost nodes, environmental and state logging, and access to a building whose configuration can be controlled. The TEBE-SENSE platform provides a suitable starting point for the sensing layer, since Section 7.7 establishes that its analytical basis is sound and the outstanding questions are exactly the field questions this programme is designed to answer [54].

The principal cost is duration rather than equipment. Drift, material interaction and seasonal variation cannot be observed in a short campaign, and it is the sustained deployment that supplies what the laboratory studies cannot. That is the element whose absence most limits what the field can currently claim, and supplying it is the contribution this programme is intended to make.

Bibliography

[1] N. Gao, Ruge Wang, Yan Wu, and Zhu Wu. "Study on impact factors of tracer gas method in investigations of gaseous pollutant transport and building ventilation". In: *Building Simulation* (2022). DOI: 10.1007/s12273-022-0947-3 (cit. on pp. 2, 10, 14, 53, 113).

[2] G. Remion, B. Moujalled, and M. E. Mankibi. "Review of tracer gas-based methods for the characterization of natural ventilation performance: Comparative analysis of their accuracy". In: *Building and Environment* (Aug. 2019). DOI: 10.1016/J.BUILDENV.2019.106180 (cit. on pp. 2, 8, 10, 12, 13, 15, 55, 61, 65, 77, 83, 92, 104).

[3] P. Lagus. "Air Leakage Measurements by the Tracer Dilution Method—A Review". In: *ASTM special technical publications* (1980), pp. 36–49. DOI: 10.1520/STP27523S (cit. on pp. 8, 65).

[4] P. Lagus and A. Persily. "A review of tracer-gas techniques for measuring airflows in buildings". In: *Ashrae Transactions* 91 (1985), pp. 1075–1087 (cit. on pp. 8, 52, 61, 77).

[5] Max H. Sherman. "Air Infiltration Measurement Techniques". In: *Lawrence Berkeley Laboratory (LBL-10705, revision)* (1998) (cit. on pp. 8, 10, 77).

[6] M. Sherman. "Tracer-gas techniques for measuring ventilation in a single zone". In: *Building and Environment* 25.4 (1990), pp. 365–374. DOI: 10.1016/0360-1323(90)90010-O (cit. on pp. 8, 12).

[7] D. Laussmann and D. Helm. "Air Change Measurements Using Tracer Gases: Methods and Results. Significance of air change for indoor air quality". In: *Chemistry, Emission Control, Radioactive Pollution and Indoor Air Quality*. InTech, July 2011. DOI: 10.5772/18600 (cit. on pp. 8, 10, 12, 13, 59).

[8] A. Persily. "Field measurement of ventilation rates". In: *Indoor air* 26 (Feb. 2016), pp. 97–111. DOI: 10.1111/ina.12193 (cit. on pp. 8, 56).

[9] A. Persily. *Tracer gas techniques for studying building air exchange*. Tech. rep. NBSIR 88-3708. National Bureau of Standards, 1988. DOI: 10.6028/NBS.IR.88-3708 (cit. on pp. 12, 13, 52).

[10] R. Dietz, T. W. D'Offavio, and R. Goodrich. "Multizone infiltration measurements in homes and buildings using a passive perfluorocarbon tracer method". In: *ASHRAE Transactions* (1985) (cit. on pp. 12, 13, 53, 93, 101).

[11] K. Cheong. "Tracer gas technology for airflow measurements in HVAC systems". In: *International Journal of Energy Research* 20.12 (Dec. 1996), pp. 1081–1093. DOI: 10.1002/(SICI)1099-114X(199612)20:12<1081::AID-ER236>3.0.CO;2-O (cit. on pp. 12, 52).

[12] J. Axley and A. Persily. *Integral mass balances and pulse injection tracer techniques*. Tech. rep. NBSIR 88-3855. National Bureau of Standards, 1988. DOI: 10.6028/NIST.IR.88-3855 (cit. on p. 12).

[13] A. Persily and W. Dols. "The Relation of CO2 Concentration to Office Building Ventilation". In: *Air Change Rate and Airtightness in Buildings*. ASTM International, 1990, pp. 77–92. DOI: 10.1520/STP17206S (cit. on pp. 12, 55).

[14] R. Almeida, E. Barreira, and P. Moreira. "A Discussion Regarding the Measurement of Ventilation Rates Using Tracer Gas and Decay Technique". In: *Infrastructures* 5.10 (Oct. 2020), p. 85. DOI: 10.3390/infrastructures5100085 (cit. on pp. 12, 55, 63, 92, 95, 110, 114, 122).

[15] Francesca Contrada, F. Causone, Yacine Allab, and A. Kindinis. "A new method for air exchange efficiency assessment including natural and mixed mode ventilation". In: *Energy and Buildings* (2021). DOI: 10.1016/j.enbuild.2021.111553 (cit. on p. 12).

[16] N. Shinohara, T. Kataoka, K. Takamine, M. Butsugan, H. Nishijima, and M. Gamo. "Modified Perfluorocarbon Tracer Method for Measuring Effective Multizone Air Exchange Rates". In: *International Journal of Environmental Research and Public Health* 7 (Aug. 2010), pp. 3348–3358. DOI: 10.3390/ijerph7093348 (cit. on pp. 12, 53, 63).

[17] S. L. Paralovo et al. "Developing a new passive tracer gas test for air change rate measurement". In: *International Journal of Ventilation* 20.3-4 (June 2020), pp. 276–286. DOI: 10.1080/14733315.2020.1777660 (cit. on pp. 14, 15, 60, 62, 63).

[18] S. L. Paralovo, M. Spruyt, J. Lauwers, M. Stranger, and J. Laverge. "The future of passive techniques for air change rate measurement". In: *39th AIVC Conference: Smart Ventilation for Buildings, Antibes Juan-Les-Pins, France*. 2018, pp. 935–942 (cit. on pp. 14, 60, 62, 114).

[19] Matthew J. Page, Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, Jennifer M. Tetzlaff, et al. "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews". In: *BMJ* 372 (2021), n71. DOI: 10.1136/bmj.n71 (cit. on pp. 17, 23).

[20] Andrew Booth, Anthea Sutton, Mark Clowes, and Marrissa Martyn-St James. *Systematic Approaches to a Successful Literature Review*. 3rd ed. London: Sage, 2021 (cit. on pp. 17, 26).

[21] Jeroen Baas, Michiel Schotten, Andrew Plume, Grégoire Côté, and Reza Karimi. "Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies". In: *Quantitative Science Studies* 1.1 (2020), pp. 377–386. DOI: 10.1162/qss_a_00019 (cit. on p. 19).

[22] Michael Gusenbauer and Neal R. Haddaway. "Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources". In: *Research Synthesis Methods* 11.2 (2020), pp. 181–217. DOI: 10.1002/jrsm.1378 (cit. on p. 19).

[23] Claes Wohlin. "Guidelines for snowballing in systematic literature studies and a replication in software engineering". In: *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering (EASE '14)*. Association for Computing Machinery, 2014, pp. 1–10. DOI: 10.1145/2601248.2601268 (cit. on p. 23).

[24] Siw Waffenschmidt, Marco Knelangen, Wiebke Sieben, Stefanie Bühn, and Dawid Pieper. "Single screening versus conventional double screening for study selection in systematic reviews: a methodological systematic review". In: *BMC Medical Research Methodology* 19.1 (2019), p. 132. DOI: 10.1186/s12874-019-0782-0 (cit. on p. 38).

[25] A. Turk, S. M. Edmonds, H. L. Mark, and G. Collins. "Sulfur hexafluoride as a gas-air tracer". In: *Environmental Science & Technology* 2.1 (1968), pp. 44–48. DOI: 10.1021/ES60013A001 (cit. on p. 52).

[26] P. Kalliokoski, R. Niemelä, and J. Salmivaara. "The tracer gas technique–a useful tool for industrial hygiene". In: *Scandinavian journal of work, environment & health* 6.2 (June 1980), pp. 123–130. DOI: 10.5271/SJWEH.2629 (cit. on pp. 52, 56, 61, 64).

[27] K. Gunn, Z. Guo, and B. Tichenor. "Tracer gas measurement of indoor-outdoor air exchange rates". In: *Proceedings of the 1994 International Symposium on Measurement of Toxic and Related Air Pollutants, Durham, North Carolina*. Air & Waste Management Association, Dec. 1994 (cit. on pp. 52, 57, 61).

[28] J. L. Barnett, M. Richard, and V. Rose. "A tracer method for quantifying contamination of building supply air: reentrainment of laboratory hood exhausts". In: *American Industrial Hygiene Association journal* 44.8 (Aug. 1983), pp. 580–582. DOI: 10.1080/15298668391405355 (cit. on p. 52).

[29] Eric J. Guffey. "Nitrous oxide as a substitute for sulfur hexafluoride in the ANSI/ASHRAE 110 Method of hood performance evaluation". Advisor: Leslie Norford. S.B. thesis. Massachusetts Institute of Technology, Department of Mechanical Engineering, 2011 (cit. on pp. 52, 56, 61).

[30] X. Yu et al. "Evaluation of nitrous oxide as a substitute for sulfur hexafluoride to reduce global warming impacts of ANSI/HPS N13.1 gaseous uniformity testing". In: *Atmospheric Environment* 176 (Mar. 2018), pp. 40–46. DOI: 10.1016/J.ATMOSENV.2017.12.015 (cit. on pp. 52, 56, 61).

[31] E. Cabaset et al. "Validation of krypton as a new tracer gas for the standardization tests of collective and individual protection systems". In: *Annals of work exposures and health* (Nov. 2023). DOI: 10.1093/annweh/wxad071 (cit. on pp. 52, 59, 62, 115).

[32] M. Bassett, C. Shaw, and R. G. Evans. "Appraisal of the sulphur hexafluoride decay technique for measuring air infiltration rates in buildings". In: *ASHRAE Transactions* 87.2 (1981) (cit. on p. 53).

[33] R. Dietz, R. Goodrich, E. Cote, and R. Wieser. "Detailed description and performance of a passive perfluorocarbon tracer system for building ventilation and air exchange measurements". In: *Measured Air Leakage of Buildings*. ASTM International, 1986, pp. 203–264. DOI: 10.1520/STP19649S (cit. on pp. 53, 61, 89).

[34] Russell N. Dietz and Edgar A. Cote. "Air infiltration measurements in a home using a convenient perfluorocarbon tracer technique". In: *Environment International* (1982). DOI: 10.1016/0160-4120(82)90059-3 (cit. on p. 53).

[35] R. Dietz, R. Goodrich, E. Cote, and R. Wieser. *Brookhaven air infiltration measurement system (BNL/AIMS) description and application*. Tech. rep. Brookhaven National Laboratory, Aug. 1983. DOI: 10.2172/5521887 (cit. on pp. 53, 61).

[36] S. Batterman, C. Jia, G. Hatzivasilis, and C. Godwin. "Simultaneous measurement of ventilation using tracer gas techniques and VOC concentrations in homes, garages and vehicles". In: *Journal of environmental monitoring : JEM* 8.2 (Feb. 2006), pp. 249–256. DOI: 10.1039/B514899E (cit. on p. 53).

[37] J. Ahmed, H. Altamirano-Medina, D. Rovas, D. Allinson, and I. Mawditt. "Design factors affecting the passive release of tracer gas for ventilation measurements". In: *E3S Web of Conferences* (2023). DOI: 10.1051/e3sconf/202339602028 (cit. on pp. 53, 61, 63).

[38] K. Ryswyk et al. “Estimation of bias with the single-zone assumption in measurement of residential air exchange using the perfluorocarbon tracer gas method”. In: *Indoor Air* 25 (Dec. 2014), pp. 610–619. DOI: 10.1111/ina.12171 (cit. on pp. 53, 61, 63).

[39] B. Leaderer, L. Schaap, and R. Dietz. “Evaluation of the perfluorocarbon tracer technique for determining infiltration rates in residences.” In: *Environmental science & technology* (1985). DOI: 10.1021/es00142a015 (cit. on p. 54).

[40] S. Batterman. “Review and Extension of CO2-Based Methods to Determine Ventilation Rates with Application to School Classrooms”. In: *International Journal of Environmental Research and Public Health* 14 (Feb. 2017). DOI: 10.3390/ijerph14020145 (cit. on pp. 55, 61, 70, 77, 81, 93, 104).

[41] C. Men, S. Wang, and Z. Zou. “Experimental study on tracer gas method for building infiltration rate measurement”. In: *Building Services Engineering Research & Technology* 41 (Mar. 2020), pp. 745–757. DOI: 10.1177/0143624420911810 (cit. on p. 55).

[42] T. Costello, N. Meador, and M. D. Shanklin. “CO2 Compared to SF6 as an Air Infiltration Tracer”. In: *Transactions of the ASAE* 27.3 (1984), pp. 844–846. DOI: 10.13031/2013.32882 (cit. on p. 55).

[43] Weihui Liang and Xiangnan Qiu. “Comparative analysis of CO2 tracer gas methods for the measurement of air change rates in student dormitory rooms with open/closed window statuses”. In: *Energy and Buildings* (2023). DOI: 10.1016/j.enbuild.2023.113363 (cit. on p. 55).

[44] G. Pei, D. Rim, S. Schiavon, and M. P. Vannucci. “Effect of sensor position on the performance of CO2-based demand controlled ventilation”. In: *Energy and Buildings* 202 (Nov. 2019), p. 109358. DOI: 10.1016/J.ENBUILD.2019.109358 (cit. on pp. 56, 63, 70, 88, 93, 99, 101, 114).

[45] D. Zhang, E. Ding, and P. Bluyssen. “Guidance to assess ventilation performance of a classroom based on CO2 monitoring”. In: *Indoor and Built Environment* 31 (Jan. 2022), pp. 1107–1126. DOI: 10.1177/1420326X211058743 (cit. on pp. 56, 61, 63, 70).

[46] C. V. M. Vouriot, M. Reeuwijk, and H. Burridge. “Robustness of point measurements of carbon dioxide concentration for the inference of ventilation rates in a wintertime classroom”. In: *Indoor Environments* (Feb. 2023). DOI: 10.1016/j.indenv.2024.100004 (cit. on pp. 56, 70).

[47] S. Laporthe, J. Virgone, and S. Castanet. “A comparative study of two tracer gases: SF6 and N2O”. In: *Building and Environment* 36.3 (Apr. 2001), pp. 313–320. DOI: 10.1016/S0360-1323(00)00005-6 (cit. on p. 56).

[48] J. Waters and M. Simons. "The Measurement of Air Infiltration in Large Single-Cell Industrial Buildings". In: *Measured Air Leakage of Buildings*. ASTM International, 1986, pp. 106–119. DOI: 10.1520/STP19644S (cit. on p. 56).

[49] M. Hori and T. Mizoguchi. "MEASUREMENT OF VENTILATION RATE USING TRACER-GAS-DECAY METHOD WITH ISOBUTENE AND PHOTO IONIZATION DETECTOR". In: *Journal of Environmental Engineering (Transactions of AIJ)* 73.626 (2008), pp. 451–455. DOI: 10.3130/AIJE.73.451 (cit. on pp. 57, 61, 64, 73, 79, 83, 85, 86, 89, 114, 117, 118).

[50] T. Tsang, K. Mui, L. Wong, K.-Y. Law, and K. Shek. "A Novel IoT-Enabled Wireless Sensor Grid for Spatial and Temporal Evaluation of Tracer Gas Dispersion". In: *Sensors (Basel, Switzerland)* 23 (Apr. 2023). DOI: 10.3390/s23083920 (cit. on pp. 58, 61, 64, 73, 79, 80, 85–87, 89, 114, 115, 117, 118, 120, 121).

[51] W. Fisk, J. Binenboym, H. Kaboli, D. Grimsrud, A. Robb, and B. J. Weber. "Multi-tracer system for measuring ventilation rates and ventilation efficiencies in large mechanically-ventilated buildings". In: *6th AIC Conference: Ventilation Strategies and Measurement Techniques, Het Meerdal Park, The Netherlands*. Paper S.5. 1985 (cit. on pp. 58, 95).

[52] D. T. Grimsrud, M. H. Sherman, J. E. Janssen, A. N. Pearman, and D. T. Harrje. "An Intercomparison of Tracer Gases Used for Air Infiltration Measurements". In: *ASHRAE Transactions* 86.1 (1980). Lawrence Berkeley Laboratory report LBL-8394. URL: https://escholarship.org/uc/item/0nn2b436 (cit. on pp. 58, 62).

[53] M. J. Jafari, S. Kalantari, R. Zendehdel, and P. Sarbakhsh. "Feasibility of Substituting Ethylene with Sulfur Hexafluoride as a Tracer Gas in Hood Performance Test by ASHRAE-110-95 Method". In: *International Journal of Occupational Hygiene* 6.1 (Oct. 2015), pp. 31–36 (cit. on pp. 58, 62).

[54] V. Gentile and M. Perino. "Low-Cost Tracer Gas Sensor for Building Ventilation". In: *Journal of Physics: Conference Series* 3140 (Nov. 2025). DOI: 10.1088/1742-6596/3140/9/092001 (cit. on pp. 58, 79, 83, 85–87, 89, 104, 110, 114, 115, 117, 118, 120, 122–124).

[55] "Room-level ventilation in schools and universities". In: *Atmospheric Environment* (2021). DOI: 10.1016/j.aeaoa.2022.100152 (cit. on p. 59).

[56] M. Samer et al. "Developing the 85Kr tracer gas technique for air exchange rate measurements in naturally ventilated animal buildings". In: *Biosystems Engineering* 109.4 (Aug. 2011), pp. 276–287. DOI: 10.1016/J.BIOSYSTEMSENG.2011.04.008 (cit. on p. 59).

[57] A. Kiwan et al. "Air exchange rate measurements in naturally ventilated dairy buildings using the tracer gas decay method with 85Kr, compared to CO2 mass balance and discharge coefficient methods". In: *Biosystems Engineering* 116.3 (Nov. 2013), pp. 286–296. DOI: 10.1016/J.BIOSYSTEMSENG.2012.11.011 (cit. on p. 59).

[58] Sarah L. Paralovo, Marianne Stranger, Borislav Lazarov, Maarten Spruyt, Joris Lauwers, Rudi Swinnen, and Jelle Laverge. "Developing a new passive tracer gas test for air change rate measurement". In: *ASHRAE / Indoor Air 2020 (16th ISIAQ Conference)* (2020) (cit. on pp. 60, 90, 114).

[59] R. Dubey et al. "Low-Cost CO2 NDIR Sensors: Performance Evaluation and Calibration Using Machine Learning Techniques". In: *Sensors (Basel, Switzerland)* 24 (Aug. 2024). DOI: 10.3390/s24175675 (cit. on pp. 61, 68, 69, 83).

[60] W. Dols and A. Persily. "A Study of Ventilation Measurement in an Office Building". In: *ASTM special technical publications* (1992). DOI: 10.1520/stp14687s (cit. on p. 62).

[61] M. R. García et al. "Review of low-cost sensors for indoor air quality: Features and applications". In: *Applied Spectroscopy Reviews* 57 (Oct. 2022), pp. 747–779. DOI: 10.1080/05704928.2022.2085734 (cit. on pp. 67–69, 77).

[62] H. Zhang and R. Srinivasan. "A Systematic Review of Air Quality Sensors, Guidelines, and Measurement Studies for Indoor Air Quality Management". In: *Sustainability* 12 (Oct. 2020), p. 9045. DOI: 10.3390/su12219045 (cit. on p. 67).

[63] S. I. Lopes et al. "Low-Cost Sensor Systems and IoT Technologies for Indoor Air Quality Monitoring: Instrumentation, Models, Implementation, and Perspectives for Validation". In: *Sensors (Basel, Switzerland)* 25 (Dec. 2025). DOI: 10.3390/s25247567 (cit. on pp. 67, 68, 74, 77).

[64] S. D. Lowther, K. Jones, X. Wang, J. D. Whyatt, O. Wild, and D. Booker. "Particulate matter measurement indoors: a review of metrics, sensors, needs and applications". In: *Environmental science & technology* (Sept. 2019). DOI: 10.1021/acs.est.9b03425 (cit. on pp. 67, 68, 71).

[65] A. Schütze et al. "Highly Sensitive and Selective VOC Sensor Systems Based on Semiconductor Gas Sensors: How to?" In: *Environments* 4.1 (Mar. 2017), p. 20. DOI: 10.3390/ENVIRONMENTS4010020 (cit. on pp. 67, 68, 72, 112).

[66] M. Verriele, D. Umba, A. Eid, S. Crunaire, N. Redon, and N. Locoge. "The Metrological Qualification of Commercial VOC Multisensor Systems: A Mandatory Step Prior to Indoor Air Quality Monitoring". In: *IEEE Sensors Journal* 24 (Apr. 2024), pp. 13153–13163. DOI: 10.1109/JSEN.2024.3368285 (cit. on pp. 67, 72).

[67] I. Demanega, I. Mujan, B. Singer, A. S. Anđelković, F. Babich, and D. Licina. "Performance assessment of low-cost environmental monitors and single sensors under variable indoor air quality and thermal conditions". In: *Building and Environment* 187 (Jan. 2021), p. 107415. DOI: 10.1016/j.buildenv.2020.107415 (cit. on pp. 67, 69, 112).

[68] Gonçalo Silva, Joana Duarte, João Santos Baptista, and João Cavaleiro Rufo. "Low-cost sensors for indoor air quality monitoring: a systematic review of accuracy, applications, and limitations". In: *Journal of the Air & Waste Management Association* 76.5 (2026), pp. 315–341. DOI: 10.1080/10962247.2026.2624795 (cit. on pp. 67, 116).

[69] Kang-Ho Ahn, Handol Lee, Hae Dong Lee, and Sang Chul Kim. "Extensive evaluation and classification of low-cost dust sensors in laboratory using a newly developed test method". In: *Indoor Air* (2019). DOI: 10.1111/ina.12615 (cit. on pp. 68, 71, 74, 86, 107, 109, 110, 113, 116, 123).

[70] T. Baur, J. Amann, C. Schultealbert, and A. Schütze. "Field Study of Metal Oxide Semiconductor Gas Sensors in Temperature Cycled Operation for Selective VOC Monitoring in Indoor Air". In: *Atmosphere* 12 (May 2021), p. 647. DOI: 10.3390/ATMOS12050647 (cit. on pp. 68, 72, 107, 110, 112, 114, 123).

[71] A. Chattopadhyay et al. "Low-Cost Formaldehyde Sensor Evaluation and Calibration in a Controlled Environment". In: *IEEE Sensors Journal* 22.12 (June 2022), pp. 11791–11802. DOI: 10.1109/jsen.2022.3172864 (cit. on pp. 68, 73).

[72] P. Kumar et al. "Real-time sensors for indoor air monitoring and challenges ahead in deploying them to urban buildings". In: *The Science of the total environment* 560–561 (Aug. 2016), pp. 150–159. DOI: 10.1016/j.scitotenv.2016.04.032 (cit. on pp. 68, 73).

[73] H. Chojer, P. Branco, F. G. Martins, M. Alvim-Ferraz, and S. I. V. Sousa. "Development of low-cost indoor air quality monitoring devices: Recent advancements". In: *The Science of the total environment* 727 (July 2020), p. 138385. DOI: 10.1016/j.scitotenv.2020.138385 (cit. on pp. 68, 74).

[74] A. Borodiņecs, A. Palcikovskis, and V. Jacnevs. "Indoor Air CO2 Sensors and Possible Uncertainties of Measurements: A Review and an Example of Practical Measurements". In: *Energies* (Sept. 2022). DOI: 10.3390/en15196961 (cit. on pp. 69, 70).

[75] A. Pastor-Fernández, A. Cerezo-Narváez, P. Montero-Gutiérrez, P. Ballesteros-Pérez, and M. Otero-Mateo. "Use of Low-Cost Devices for the Control and Monitoring of CO2 Concentration in Existing Buildings after the COVID Era". In: *Applied Sciences* 12.8 (Apr. 2022), p. 3927. DOI: 10.3390/app12083927 (cit. on pp. 69, 80, 82, 85, 88, 89).

[76] David Honan, John Garvey, J. Littlewood, Matthew Horrigan, and John Gallagher. "Low-Cost CO2 Sensors: On-Site Performance Evaluation and Co-Location Correction Procedure for Reliable Ventilation Assessments in Schools". In: *Sensors (Basel, Switzerland)* (2026). DOI: 10.3390/s26041265 (cit. on pp. 70, 106, 108).

[77] Daniel Marín, Alba Alegría-Sala, Lluc Canals Casals, Marcel Macarulla, and Jordi Fonollosa. "The reliability of CO_2 measurements using low-cost sensors: a study of sensor positioning and ventilation strategies in classrooms". In: *Indoor Air* 2025 (2025), p. 5517242. DOI: 10.1155/ina/5517242 (cit. on pp. 70, 88).

[78] N. Mahyuddin and H. Awbi. "A Review of CO2 Measurement Procedures in Ventilation Research". In: *International Journal of Ventilation* 10 (Mar. 2012), pp. 353–370. DOI: 10.1080/14733315.2012.11683961 (cit. on p. 70).

[79] J. Li, S. Mattewal, S. Patel, and P. Biswas. "Evaluation of Nine Low-cost-sensor-based Particulate Matter Monitors". In: *Aerosol and Air Quality Research* 20.2 (Feb. 2020), pp. 254–270. DOI: 10.4209/aaqr.2018.12.0485 (cit. on p. 71).

[80] Z. Wang, W. Delp, and B. Singer. "Performance of low-cost indoor air quality monitors for PM2.5 and PM10 from residential sources". In: *Building and Environment* 171 (Mar. 2020), p. 106654. DOI: 10.1016/j.buildenv.2020.106654 (cit. on p. 71).

[81] Ruikang He, Taewon Han, Daniel Bachman, Dominick J. Carluccio, Rudolph Jaeger, Jie Zhang, Sanjeevi Thirumurugesan, Clinton Andrews, and Gediminas Mainelis. "Evaluation of two low-cost PM monitors under different laboratory and indoor conditions". In: *Aerosol Science and Technology* 55.3 (2021), pp. 316–331. DOI: 10.1080/02786826.2020.1851649 (cit. on p. 71).

[82] Zikang Feng, Lina Zheng, and Dou Liu. "Analytical performance and calibration strategies of low-cost particulate matter sensors for indoor and workplace monitoring: a review". In: *Analytical Methods* 18 (2026), pp. 39–52. DOI: 10.1039/D5AY01554E (cit. on pp. 71, 106, 111).

[83] Jakub Kolarik, Nadja Lynge Lyng, Rossana Bossi, Rongling Li, Thomas Witterseh, Kevin Michael Smith, and Pawel Wargocki. "Application of cluster analysis to examine the performance of low-cost volatile organic compound sensors". In: *Buildings* 13.8 (2023), p. 2070. DOI: 10.3390/buildings13082070 (cit. on p. 72).

[84] J. Saini, M. Dutta, and G. Marques. "Indoor Air Quality Monitoring Systems Based on Internet of Things: A Systematic Review". In: *International Journal of Environmental Research and Public Health* 17 (July 2020). DOI: 10.3390/ijerph17144942 (cit. on p. 74).

[85] H. Zhang, R. Srinivasan, and V. Ganesan. "Low Cost, Multi-Pollutant Sensing System Using Raspberry Pi for Indoor Air Quality Monitoring". In: *Sustainability* 13.1 (Jan. 2021), p. 370. DOI: 10.3390/su13010370 (cit. on p. 74).

[86] Hanin Othman, R. Azari, and Tamy Guimarães. "Low-Cost IoT-based Indoor Air Quality Monitoring". In: *Technology|Architecture + Design* (2024). DOI: 10.1080/24751448.2024.2405403 (cit. on p. 74).

[87] J. Saini, M. Dutta, and Gonçalo Marques. "Sensors for indoor air quality monitoring and assessment through Internet of Things: a systematic review". In: *Environmental Monitoring and Assessment* (2021). DOI: 10.1007/s10661-020-08781-6 (cit. on p. 74).

[88] A. Rusch and T. Rösgen. "An Internet of Things Sensor Array for Spatially and Temporally Resolved Indoor Climate Measurements". In: *Sensors (Basel, Switzerland)* 22 (June 2022). DOI: 10.3390/s22124377 (cit. on p. 74).

[89] S. Manu and A. Rysanek. "A Co-Location Study of 87 Low-Cost Environmental Monitors: Assessing Outliers, Variability, and Uncertainty". In: *Buildings* (Sept. 2024). DOI: 10.3390/buildings14092965 (cit. on pp. 74, 109, 110, 123).

[90] N. Zaky, T. Li, and H. Stopps. "Validating the performance of low-cost IAQ sensors through co-location". In: *Journal of Building Physics* 49 (Oct. 2025), pp. 829–851. DOI: 10.1177/17442591251367436 (cit. on pp. 74, 86, 107, 110, 115, 124).

[91] Olivier Ramalho, Ingrid Schneider, Caroline Marchand, and Corinne Mandin. "Low-cost sensors for indoor air quality: performance assessment in laboratory conditions". In: *Indoor Air 2020 (16th ISIAQ Conference)* (2020) (cit. on pp. 74, 86, 109, 116).

[92] Caroline Marchand, Amandine Fievet, Olivier Ramalho, Jessica Queron, and Laurent Spinelle. "Low-cost sensors for indoor air quality: performance assessment in simulated real conditions". In: *Indoor Air 2020 (16th ISIAQ Conference)* (2020) (cit. on pp. 74, 86, 109).

[93] M. Benedettelli, B. Naticchia, A. Carbonari, and M. Pascucci. "Testing of a Tracer Gas Based Measurement Procedure to Assess Air Change Rates in Buildings". In: *Proceedings of the 32nd International Symposium on Automation and Robotics in Construction (ISARC)*. IAARC, June 2015, pp. 201–207. DOI: 10.22260/ISARC2015/0025 (cit. on pp. 79, 81, 85, 88, 89, 115).

[94] A. O. Pérez, B. Bierer, L. Scholz, J. Wöllenstein, and S. Palzer. "A Wireless Gas Sensor Network to Monitor Indoor Environmental Quality in Schools". In: *Sensors (Basel, Switzerland)* 18 (Dec. 2018). DOI: 10.3390/s18124345 (cit. on pp. 79, 82, 85, 89, 93, 101, 115, 117, 121).

[95] Raja Singh and Anil Dewan. "Low cost tool for finding ventilation rates for airborne infection control in the built environment". In: *Indoor Air 2020 (16th ISIAQ Conference)* (2020) (cit. on p. 82).

[96] J. Coleman and F. Meggers. "Sensing of Indoor Air Quality—Characterization of Spatial and Temporal Pollutant Evolution Through Distributed Sensing". In: *Frontiers in Built Environment* 4 (May 2018). DOI: 10.3389/fbuil.2018.00028 (cit. on p. 85).

[97] Aniebietabasi Ackley, Michael Donn, Geoff Thomas, Wallace Enegbuma, and Shuva Chowdhury. "Use of carbon dioxide (CO_2) monitors to assess ventilation effectiveness in schools". In: *Journal of Sustainable Architecture and Civil Engineering* 32.1 (2023), pp. 130–144. DOI: 10.5755/j01.sace.32.1.31536 (cit. on p. 88).

[98] Brodie W. Hobson, Andre A. Markus, H. Burak Gunay, and Farzeen Rizvi. "Minimum sensor grid density and configuration to enable CO2-based demand-controlled ventilation in an office building". In: *Energy and Buildings* (2023). DOI: 10.1016/j.enbuild.2023.113536 (cit. on p. 88).

[99] Sarah L. Paralovo, M. Stranger, B. Lazarov, M. Spruyt, and J. Laverge. "Adapted tracer gas test for passive measurement of total air change rates using alternative tracer substance". In: *Science and Technology for the Built Environment* (2021). DOI: 10.1080/23744731.2021.1997061 (cit. on pp. 89, 90).

[100] Tapan Patel and Constandinos Mitsingas. "Comparison of Blower Door and Tracer Gas Testing Methods for Determination of Air Infiltration Rates through Building Envelopes at Normal Operating Conditions". In: *ASME International Conference on Energy Sustainability (ES2011)* (2011) (cit. on pp. 92, 94, 101, 115, 117, 118, 122).

[101] Alan Vega Pasos et al. “Experimental study on the measurement of Building Infiltration and Air Leakage rates (at 4 and 50 Pa) by means of Tracer Gas methods, Blower Door and the novel Pulse technique in a Detached UK Home”. In: *AIVC Conference Proceedings* (2018) (cit. on pp. 92, 94, 101, 117, 118).

[102] K. W. Cheong. “Airflow measurements for balancing of air distribution system — tracer-gas technique as an alternative?” In: *Building and Environment* (2001) (cit. on pp. 92, 93, 97, 101, 118, 122, 123).

[103] Michael D. Sohn, Pamela Reynolds, Navtej Singh, and Ashok J. Gadgil. “Rapidly Locating and Characterizing Pollutant Releases in Buildings”. In: *Journal of the Air & Waste Management Association* (2002). DOI: 10.1080/10473289.2002.10470869 (cit. on pp. 92, 98, 101, 117, 123).

[104] Hao Cai, Xianting Li, Zhilong Chen, and Mingyang Wang. “Rapid identification of multiple constantly-released contaminant sources in indoor environments with unknown release time”. In: *Building and Environment* (2014). DOI: 10.1016/j.buildenv.2014.06.006 (cit. on pp. 92, 98, 101, 117).

[105] Vladimir Vukovic, Paulo Cesar Tabares-Velasco, and Jelena Srebric. “Real-Time Identification of Indoor Pollutant Source Positions Based on Neural Network Locator of Contaminant Sources and Optimized Sensor Networks”. In: *Journal of the Air & Waste Management Association* (2010). DOI: 10.3155/1047-3289.60.9.1034 (cit. on pp. 92, 99, 101, 117, 122, 123).

[106] James C. Matthews, M. Anwar H. Khan, Matthew D. Wright, Prem K. Perumal, Carl J. Percival, Ian D. Bull, Ian J. Craddock, and Dudley E. Shallcross. “Indoor Exchange Rates and Penetration From Outdoors in an Instrumented Terraced House (Townhouse) Using Gas Tracers: Implications for Particles and Gases Indoors”. In: *Indoor Air* (2024). DOI: 10.1155/2024/9204433 (cit. on p. 93).

[107] William P. Bahnfleth, Grenville K. Yuill, and Brian W. Lee. “Protocol for Field Testing of Tall Buildings to Determine Envelope Air Leakage Rate”. In: *ASHRAE Transactions (RP-935)* (1999) (cit. on pp. 93, 95, 101, 115).

[108] Stuart Batterman, Fu-Chen Su, Adam Wald, Farhan Watkins, Christopher Godwin, and Geoffrey Thun. “Ventilation rates in recently constructed U.S. school classrooms”. In: *Indoor Air* 27.5 (2017), pp. 880–890. DOI: 10.1111/ina.12384 (cit. on p. 93).

[109] Bart Merema, Muhannad Delwati, Maarten Sourbron, and Hilde Breesch. “Demand controlled ventilation (DCV) in school and office buildings: lessons learnt from case studies”. In: *Energy and Buildings* 172 (2018), pp. 349–360. DOI: 10.1016/j.enbuild.2018.04.065 (cit. on p. 93).

[110] W. Fisk, R. Prill, and O. Seppanen. "Commercial building ventilation measurements using multiple tracer gases". In: *9th AIVC Conference: Effective Ventilation, Gent, Belgium*. Paper 9. 1988 (cit. on p. 96).

[111] J. Reardon, C. Shaw, and F. Vaculik. "Air change rates and carbon dioxide concentrations in a high-rise office building". In: *ASHRAE Transactions* (1994) (cit. on p. 96).

[112] Christina Higgins, Prashant Kumar, and Lidia Morawska. "Indoor air quality monitoring and source apportionment using low-cost sensors". In: *Environmental Research Communications* (2024). DOI: 10.1088/2515-7620/ad1cad (cit. on pp. 99, 116).

[113] Jinze Wang, Wei Du, Yali Lei, Yuanchen Chen, Zhenglu Wang, Kang Mao, Shu Tao, and Bo Pan. "Quantifying the dynamic characteristics of indoor air pollution using real-time sensors: current status and future implication". In: *Environment International* 175 (2023), p. 107934. DOI: 10.1016/j.envint.2023.107934 (cit. on p. 100).

[114] Shelly L. Miller, K. Leiserson, and W. Nazaroff. "Nonlinear Least-Squares Minimization Applied to Tracer Gas Decay for Determining Airflow Rates in a Two-Zone Building". In: *Indoor Air* (1997). DOI: 10.1111/j.1600-0668.1997.t01-1-00008.x (cit. on p. 105).

[115] T. D'ottavio, G. Senum, and R. Dietz. "Error analysis techniques for perfluorocarbon tracer derived multizone ventilation rates". In: *Building and Environment* (1988). DOI: 10.1016/0360-1323(88)90003-0 (cit. on pp. 105, 110, 113).

[116] M. Sherman. "Analysis of errors associated with passive ventilation measurement techniques". In: *Building and Environment* (1989). DOI: 10.1016/0360-1323(89)90002-4 (cit. on pp. 106, 110, 114).

[117] J. Tryner, Mollie Phillips, C. Quinn, Gabe Neymark, A. Wilson, S. Jathar, Ellison Carter, and J. Volckens. "Design and Testing of a Low-Cost Sensor and Sampling Platform for Indoor Air Quality". In: *Building and environment* (2021). DOI: 10.1016/j.buildenv.2021.108398 (cit. on p. 108).

[118] Alejandro Moreno-Rangel, Tim Sharpe, Filbert Musau, and Gráinne McGill. "Field evaluation of a low-cost indoor air quality monitor to quantify exposure to pollutants in residential environments". In: *Journal of Sensors and Sensor Systems* 7.2 (2018), pp. 373–388. DOI: 10.5194/jsss-7-373-2018 (cit. on p. 109).

[119] Iain Walker, Woody Delp, and Brett Singer. "Are low-cost sensors good enough for IAQ controls?" In: *REHVA Journal* 56.1 (Nov. 2019). Based on a paper presented at the 39th AIVC, 7th TightVent and 5th venticool Conference, Smart Ventilation for Buildings, Antibes Juan-Les-Pins, France, 18–19 September 2018 (cit. on p. 109).