15TH IEA HEAT PUMP CONFERENCE
MAY 26–29, 2026 | HOFBURG
VIENNA 2026
DECARBONISATION THROUGH INNOVATION

"Performance gaps and quality assurance challenges in the UK heat pump roll-out: evidence from 1,700 metered installations"

Colin Meek[a], Jamie Newman[a]

[a]**rb&m**, Hartfield Rd, London, SW19 3TH, UK

Abstract

This paper presents a comprehensive analysis of in-situ performance data from 1,700 domestic heat pump installations carried out under the UK's Renewable Heat Incentive (RHI) between 2017 and 2022. The dataset – obtained by UK's Department for Energy Security and Net Zero (DESNZ) – includes approximately 300 ground-source systems, providing a unique large-scale insight into ground source heat pump (GSHP) performance. The analysis compares measured *Seasonal Performance Factors (SPF)* with installer design forecasts based on the *Seasonal Coefficient of Performance (SCOP)* metric.
Although a substantial proportion of heat pumps met the SPF ≥ 3.0 benchmark (30% of ASHPs and 67% of GSHPs), median SPFs fell notably short of installer design forecasts: 2.69 for ASHPs (design SCOP 3.59) and 3.26 for GSHPs (design SCOP 3.93), representing performance gaps of 0.97 and 0.70 respectively. The correlation between measured and forecast efficiencies was weak (Pearson r = 0.225 for ASHPs; r = 0.228 for GSHPs). These findings closely align with results from the independent Electrification of Heat (EoH) Demonstration Project, which reported a comparable performance gap (ASHP SPF_{H4} difference = 0.85), validating both studies. Further analysis of the dataset from the EoH project is described in this paper.
The study finds that although heat-pump technology is reliable, many installations underperform, risking consumer trust and hindering UK decarbonisation. It recommends evolving regulations to focus on outcome-based assurance – such as public performance databases, post-installation monitoring incentives and the establishment of minimum in-situ performance thresholds (e.g. SPF ≥ 2.5) or contractual guarantees assigning outcome responsibility to installers. These strategies aim to de-risk the market for consumers, discourage "fit-and-forget" installation practices and ensure transparency. They also reflect the principles of regulatory science, emphasising evidence-based, adaptive policymaking informed by measured real-world outcomes. Building trust and reducing financial risk are crucial for broader heat pump adoption over gas boilers.

Keywords: heat pumps, heat pump performance, ASHPs, GSHPs, heat pump monitoring, heat pump optimisation, heat pump efficiency

1. Introduction

The UK's Climate Change Committee (CCC) estimates that 19 million heat pump installations are essential to the country's decarbonisation targets [1]. Although installation rates are improving, the CCC has said that the implementation of electrification of heat using heat pumps is off-track and that the UK lags far behind many other European countries [2]. The Committee has also repeatedly stressed that consumer confidence is critical to the successful rollout of the UK's strategy for heat pumps and has linked the pace of change directly with consumer trust [3].

The pace of heat pump deployment in the UK is closely tied to the relative price of electricity compared to gas and consumer perceptions about future running costs. Because domestic electricity prices bear a disproportionate share of policy and network levies and are further influenced by marginal pricing in wholesale power markets, they remain around three to four times higher per kWh than gas, weakening the economic case for heat pumps. Both the CCC and the government acknowledge that consumer confidence depends on whether a heat pump will deliver lower or at least comparable heating costs to a gas boiler [4].

Previous studies on heat pump performance in the UK have reported disappointing results. For example, a UK Government funded study on the performance of heat pumps deployed as part of the Renewable Heat Premium Payment (RHPP) incentive scheme found that the median seasonal performance factor (SPF) for AHPs in the monitored 292 installations was **2.65** (SPF_{H2}) and **2.44** (SPF_{H4}), reflecting substantial variation in real-world performance. The data was collected from 2013 to 2015 and published in 2017 [5].

Seasonal Performance Factors (SPFs) are most commonly calculated using system boundaries defined by the EU-funded SEPEMO (Seasonal Performance factor and Monitoring) project, which established standardised boundaries for heat-pump field trials: the H2 boundary captures the electrical input to the heat-pump unit itself, while the H4 boundary extends this to include all space-heating related electricity consumption, such as circulation pumps and controls, thereby representing whole-system performance.

This paper uses data from 1,700 heat pump installations (including nearly 300 GSHPs) to test whether installer performance forecasts (based on SCOP) are reliable predictors of real-world system performance. The installations are subject to mandatory metering by the Office of Gas and Electricity Markets (Ofgem) under the UK's Renewable Heat Incentive (RHI) scheme, with the data for this study drawn from monitoring of installations carried out between October 2017 and December 2022.

The results are contextualized by comparison with the EoH Demonstration Project (2024) [6]. The EoH project was funded by the UK’s DESNZ and designed to assess the feasibility of large-scale domestic heat-pump deployment across Great Britain. The project was delivered through a structured consortium model: Energy Systems Catapult (ESC) led as Management Contractor (with LCP Delta and Oxford Computer Consultants), supported by three regional Delivery Contractors – E.ON, OVO Energy, and Warmworks Scotland – responsible for recruitment, home surveys, system design, installation, and ongoing monitoring. Metering equipment was used to record electricity input and heat output and evaluate real-world heat-pump performance. This paper also includes further analysis of that publicly available EoH dataset. The strong alignment observed between these independent studies enhances confidence in both the outcomes and their potential policy implications.

2. Background

2.1 Previous Research

In 2019, the author obtained a large Ofgem dataset under the Environmental Information Regulations containing anonymised quarterly metered heat output and electricity consumption for domestic air-source and ground-source heat pumps participating in the RHI’s metering-for-payment scheme. As the dataset also included the installers’ compulsory Microgeneration Certification Scheme (MCS) performance estimates, it was possible to compare *measured in-situ* efficiency (SPF) with the *design performance forecasts* provided prior to installation to Ofgem. After filtering for systems with at least one full year of valid data, nearly 600 installations were available for analysis, though relatively few were installed *after* the introduction of the revised UK installation standard: the MCS MIS 3005 in late 2017 [7]. That new standard implemented a modified SCOP product metric compulsory for all performance estimates. The final analysis, published in 2022 [8], reported a significant ‘performance gap’ between the SPF results and the design performance forecasts provided to consumers by installers.

2.2 Performance estimates based on the SCOP Metric

The European standards EN 14825, EN 16147, and EN 14511 underpin the testing and rating of heat pump performance within the EU’s *Ecodesign* (Regulations 813/2013 and 814/2013) and Energy Labelling (Regulations 811/2013 and 812/2013) frameworks. Together they define how manufacturers must determine and declare the SCOP and related efficiency values displayed on consumer energy labels. Intended to ensure comparability and transparency, these harmonised standards express efficiency under reproducible test conditions and seek to support informed purchasing decisions. In the UK, consumer-facing information on heat-pump performance is dominated by references to the product SCOP metric.

Reference to the SCOP metric became pervasive when, in October 2017, the MCS adopted it as the formal basis for installers’ consumer performance estimates under MIS 3005 [7]. Installers were required to provide a design SCOP forecast using a standardised template document. In March 2025, MCS withdrew the SCOP methodology as the compulsory basis for those estimates and introduced a new algorithm designed to predict

Seasonal Performance Factors (SPFs) using parameters such as target flow temperature and system configuration [9].

Nevertheless, SCOP values continue to underpin many promotional claims and remain embedded in consumer information tools – including online calculators that estimate potential running-cost savings (for example, Project Ambient [10]).

As O'Hegarty et al. observe: "These specified product performance ratings are a means for the architect, heating specifier, and the consumer to discern the optimum efficient technology and to estimate use and lifecycle costs. But if the product performance ratings are unrepresentative of the actual field/in-use performance their significance is lost, and more crucially the consumer may be misdirected" [11]. This highlights the risk that SCOP metrics, though valuable for laboratory benchmarking, can mislead if used to predict real-world outcomes.

The UK Competition and Markets Authority (2023) has similarly cautioned that where businesses advertise maximum performance achievable only under specific laboratory conditions that do not reflect most consumers' circumstances, such claims may constitute misleading marketing [12].

The analysis presented here provides one of the few truly large-scale comparisons between measured in-situ seasonal efficiencies and the SCOP-based estimates supplied at installation. While both datasets have recognised limitations, their consistency across more than 2,000 installations provides robust evidence of a persistent performance gap with significant implications for policy design, installer accreditation, and consumer confidence in the heat-pump market.

2. Methodology

The Ofgem data set contained anonymised information for heat pump systems installed under the RHI scheme and subject to 'metering for payment' (bi-valent systems or second homes). Strict quality criteria were deployed to make sure that the meter readings for all installations started after the MCS installation standard MIS 3005 V5 became compulsory in October 2017. Additionally, installations were only included where there was at least one whole year of contiguous clean data. Rigorous data cleaning identified and removed erroneous readings.

Ofgem rules relating to 'metering for payment' installations specified that 'all electrical input to the system' be metered to provide electricity consumption. This approach approximates the SEPEMO H4 boundary [13]; however, as the rules did not explicitly reference the SEPEMO boundary definitions, the SPF results derived from the Ofgem dataset are labelled $Ofgem_{H4}$ rather than SPF_{H4}. The other project limitations are discussed in Section 4. The calculated in-situ SPF in this analysis is compared directly to the SCOP-based design forecast. In other words, the actual consumer experience is compared to the installer system performance prediction provided at the time.

SPFs were calculated as the total metered heat output divided by the total electricity input over a minimum of one full year of operation. Paired t-tests were used to assess the statistical significance of the performance gap between SCOP estimates and measured SPF values, and Pearson correlation coefficients were calculated to examine the strength of the SCOP-SPF relationship. Frequency distributions are reported in 0.25 coefficient increments. The final analytical sample comprised 1,431 ASHPs and 286 GSHPs, and the resulting performance distributions were compared with the published EoH field trial data for validation.

3. Results

The results demonstrate that a significant proportion of heat pumps were performing well:

- 67% GSHPs and 30% ASHPs achieved SPF ≥3.0;
- 33% GSHPs and 8% ASHPs achieved SPF ≥3.5;
- 15% GSHPs achieved SPF ≥4.0; and
- nearly 6% ASHPs and 16% GSHPs met or exceeded the design SCOP.

These results support the EoH assessment that heat pumps can operate with good efficiencies and are an important low-carbon heating technology across all UK housing archetypes [6].

The median SPFs calculated were, however, substantially lower than the design SCOPs (Table 1). The median SPF ($Ofgem_{H4}$) for the whole ASHP sample was 2.69 and the median SPF ($Ofgem_{H4}$) for GSHPs was 3.26. Table 1 also provides the ASHP results from the EoH research.

Table 1. Median and Mean Measured Efficiencies and Median and Mean Design Forecasts

Dataset	Median Actual Efficiency – SPF [IQR]	Mean Actual Efficiency – SPF [95% CI]	Median Design Forecast Efficiency – SCOP [IQR]	Mean Design Forecast Efficiency – SCOP [95% CI]
Ofgem Data ASHPs (1431) $Ofgem_{H4}$	**2.69** [2.26, 3.07]	2.65 [2.61, 2.68]	**3.59** [3.41, 3.86]	3.61 [3.59, 3.63]
Ofgem Data GSHPs (286) $Ofgem_{H4}$	**3.26** [2.83, 3.64]	3.24 [3.16, 3.32]	**3.93** [3.59, 4.29]	3.95 [3.90, 4.00]
Ofgem Data ASHPs (2022 installs only)(435) $Ofgem_{H4}$	**2.74** [2.26, 3.11]	2.67 [2.61, 2.73]	**3.67** [3.46, 3.92]	3.71 [3.68, 3.74]
Ofgem Data GSHPs (2022 installs only)(116) $Ofgem_{H4}$	**3.34** [2.88, 3.73]	3.31 [3.18, 3.43]	**3.99** [3.73, 4.43]	4.06 [3.97, 4.14]
EoH Data ASHPs (428) SPF_{H4}	**2.78** [2.55, 3.05]	2.81 [2.76,2.85]	**3.65**	3.65
EoH Data ASHPs (428) SPF_{H2}	**2.93** [2.67, 3.19]	2.95 [2.90, 2.99]	〃	〃

Note: $Ofgem_{H4}$ approximates to SEPEMO SPF_{H4} (see Section 2).

3.1. Validation: Comparison with EoH Study

The Ofgem dataset and EoH field trial results closely align. As shown in Table 2, both studies found a comparable proportion of installations where the calculated SPFs were lower than the design SCOPs. Specifically, the Ofgem data indicated that the mean ASHP SPF ($Ofgem_{H4}$) was 27% below the design SCOP, while the EoH study observed a similar relative difference of 23% (SPF_{H4}). For the 2022 cohort in the Ofgem analysis, the relative difference was 28%. Furthermore, Table 2 demonstrates that the EoH study identified a significant performance gap between the SPF_{H2} results and the design SCOPs, suggesting that the underlying causes of this gap extend well beyond purely technical differences between the H2 and H4 boundaries.

Table 2. Comparing actual performance (SPF) with design forecasts (SCOP)

Dataset	% of installs SPF lower than SCOP	Mean difference between SPF and SCOP	Relative difference (%) between SPF and SCOP
Ofgem Data ASHPs (1431) $Ofgem_{H4}$	**94.4**	**0.97**	**26.8**
Ofgem Data GSHPs (286) $Ofgem_{H4}$	**83.7**	**0.70**	**17.9**
Ofgem Data ASHPs (2022 installs only)(435) $Ofgem_{H4}$	**95.0**	**1.04**	**28.1**
Ofgem Data GSHPs (2022 installs only)(116) $Ofgem_{H4}$	**84.5**	**0.75**	**18.5**
EoH Data ASHPs (413) SPF_{H4}	**92.3**	**0.85**	**23.2**
EoH Data ASHPs (413) SPF_{H2}	**90.99**	**0.72**	**19.2**

Figure 1 presents the complete ASHP sample frequency distribution in increments of 0.25 coefficients based on the Ofgem data analysis. Additionally, an examination of the publicly accessible EoH dataset for this study set out in Figure 2 provides a comparable distribution for reference. Both show very wide distribution of performance and poor alignment with the SCOP design forecasts. Whilst both show similar clustering between SPF 2.5 and 3.0, a substantial minority of installations perform with SPFs below 2.5: 38% of ASHPs in the analysis of Ofgem data and 22% of those in the EoH field trial. At the other end of the spectrum, 8% of those in the Ofgem data analysis performed with an SPF above 3.5, while the equivalent finding from the installs in the EoH trial was 6%. Overall, the tighter distribution evident for the EoH sample may reflect the very small number of installation contractors used in that study (three). Whilst the Ofgem dataset did not identify individual installation companies, the wide geographical distribution of installations throughout Great Britain indicates that several hundred installers were likely involved.

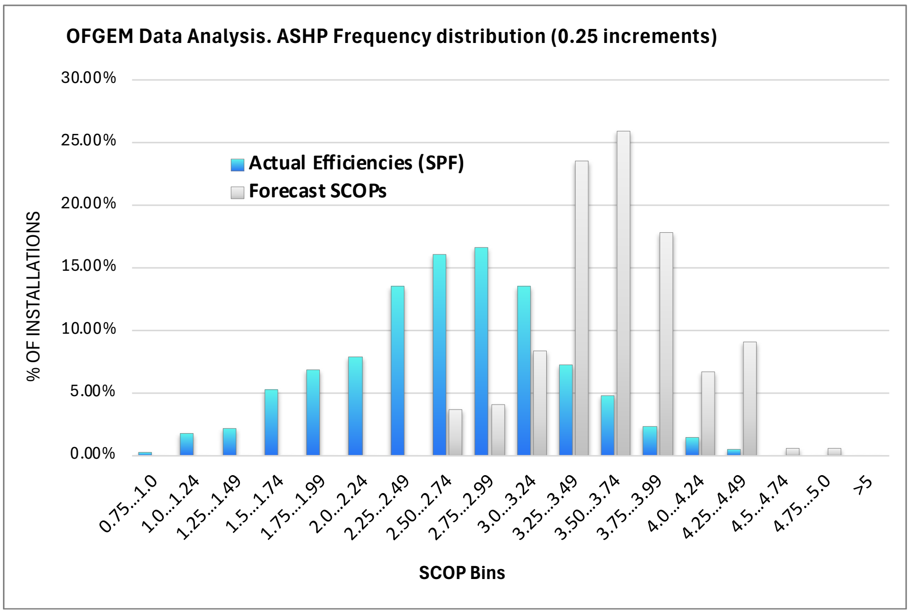

Fig 1. Frequency distribution of ASHP SPFs (Ofgem$_{H4}$) compared to installer forecast SCOPs. Sample size 1431.

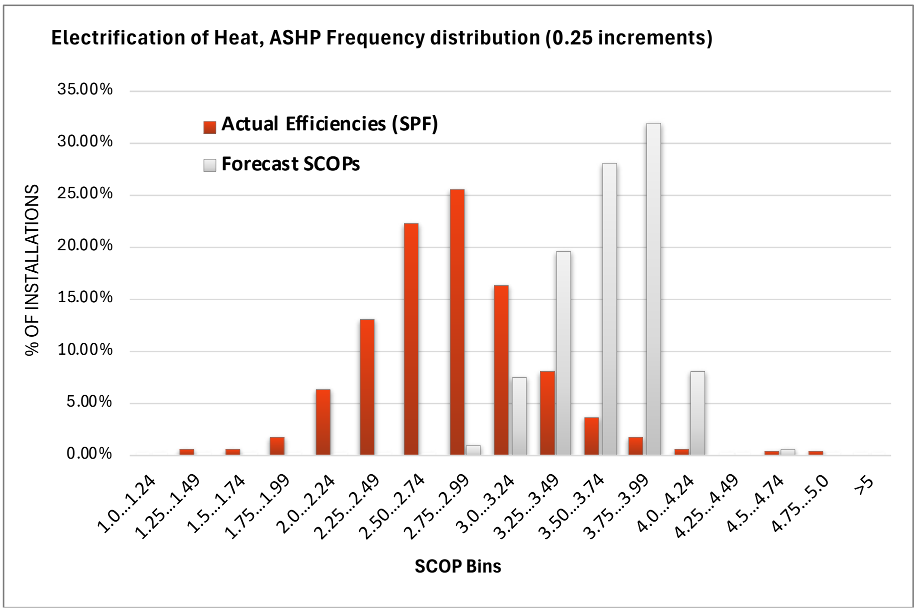

Fig 2. Frequency distribution of ASHP SPF$_{H4}$ Electrification of Heat Dataset compared to installer design SCOPs. Sample size 428.

3.2. Lack of Correlation between SPFs and design SCOPs

Figures 3 and 4 provide visualisations of the correlation between the SPFs (Ofgem$_{H4}$) and the design SCOP. Both show that relatively high SCOPs are almost never achieved in-situ whilst only a small proportion of SPFs match or surpass the design forecasts. The statistical analysis found: Pearson r = 0.225 (ASHPs) and r = 0.228 (GSHPs) – indicating weak correlation for both samples. The authors of the EoH field trial report also concluded that there was “little to no correlation” between the designed SCOPs and SPFs in that analysis [6].

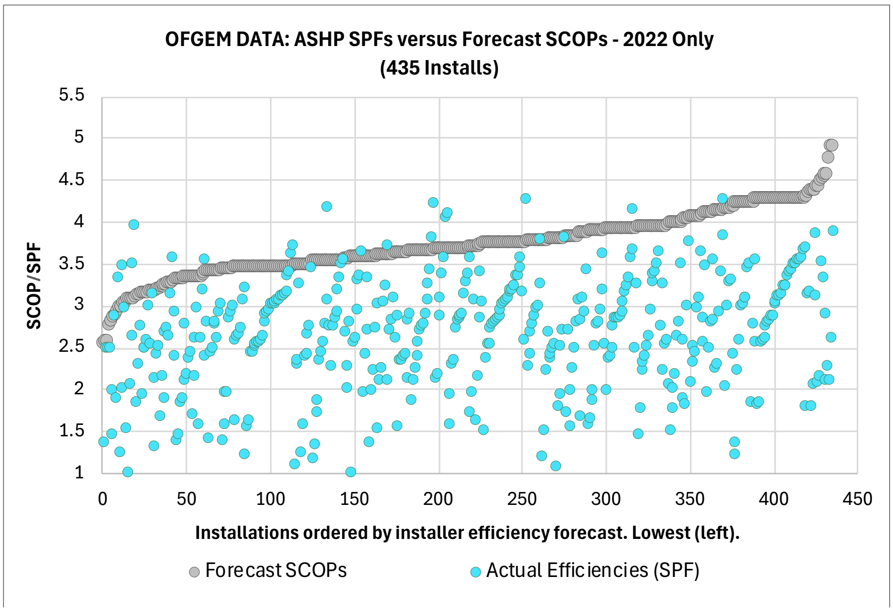

Fig 3. Ofgem Data. 2022 ASHP installation cohort. Installation SCOPs (grey) with corresponding blue markers showing SPFs ($Ofgem_{H4}$). Sample size 435.

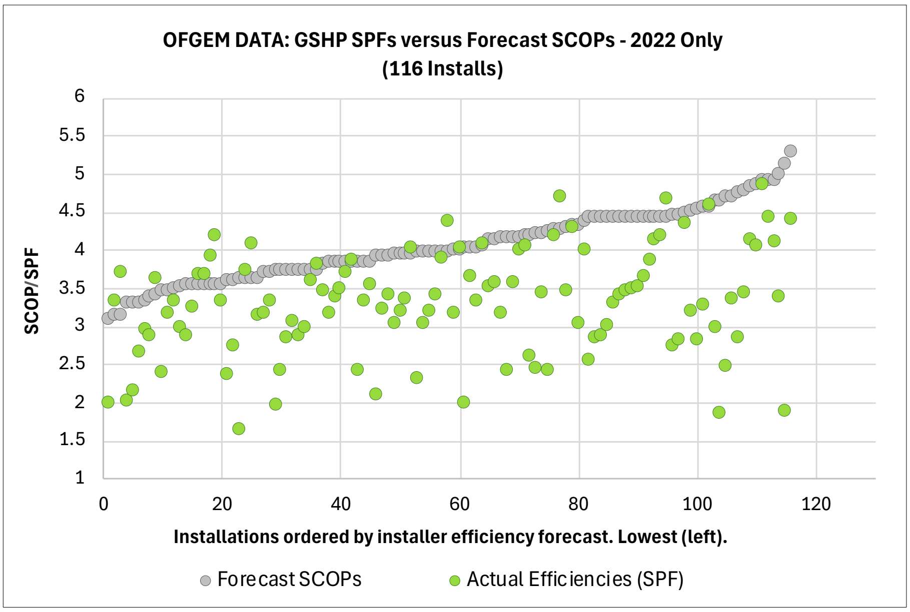

Fig 4. Ofgem Data. 2022 GSHP installation cohort. Installation SCOPs (grey) with corresponding green markers showing SPFs ($Ofgem_{H4}$). Sample size 116.

3.3. Performance Trends Over Time

The SPFs calculated in the analysis of Ofgem data and in the EoH study are significantly better than those reported in the study of Renewable Heat Premium Payment (RHPP) installations published in 2017 [5]. The median SPF_{H4} in that study of 292 heat pump systems was 2.44.

However, the analysis of Ofgem data by installation cohort reveals divergent trends (Figure 5). GSHP in-situ performance improved from 3.07 reported in the original analysis of Ofgem data [8] to 3.34 for the 2022 cohort of installations ($Ofgem_{H4}$). The performance gap for GSHPs during that period remained relatively stable. Conversely, the in-situ performance for ASHPs showed no significant improvement, while the mean design SCOP rose from 3.3 to 3.71, resulting in a widening performance gap. Table 1 shows that the mean difference between SPF ($Ofgem_{H4}$) and SCOP for the 2022 cohort of installations was 1.04. This suggests increasingly optimistic forecasts without corresponding real-world improvement. There is no clear explanation for the absence of observable performance improvement in ASHPs over the monitored period; the potential reasons for the persistence of the performance gap are therefore explored in Section 5.2.

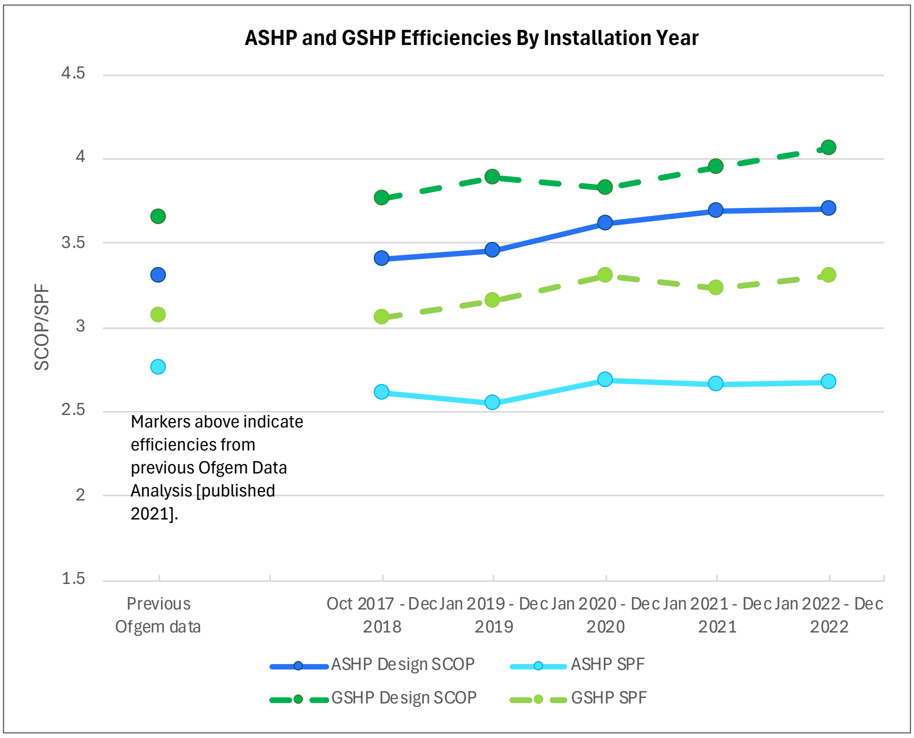

Fig 5. Trends over time. The Ofgem analysis by installation cohort.

3.4. Validation

Despite different data sources and methodologies, both the Ofgem data analysis and the EoH study identify:

- similar median ASHP SPF (2.69 vs 2.78 – only 0.09 difference);
- a similar performance gap (0.97 vs 0.85); and
- weak correlation between SCOP and SPF.

This alignment enhances confidence that the observed performance gap and weak correlation are consistent, systematic results rather than anomalies attributable to specific samples.

4. Limitations

The limitations related to the electricity metered boundary are described in Section 2. The analysis of the Ofgem data is based solely on installations subject to *Metering for Payment* under the RHI, a subset that is atypical – often second homes or properties with hybrid or bi-valent systems. As such, the results may not represent the wider UK housing stock or the general heat-pump market. The dataset comprises metered electricity and heat-output data but lacks full system-monitoring detail such as internal temperatures, flow temperatures, domestic-hot-water demand, or supplementary-heat contributions.

It is important to interpret these limitations within their proper context. The Ofgem dataset offers a reliable account of heat-pump electricity consumption and heat generation, given the requirements for correct installation of metering equipment and the obligation of accurate reading submissions under the RHI scheme. The data-collection process is designed to minimize unit and reporting errors, and this study has implemented rigorous quality controls, that excluded anomalies and retained only those installations with at least a full year of contiguous, high-quality data. Despite the secondary nature of the data, the analysis covers 1,700 heat-pump installations, permitting meaningful comparisons across annual cohorts. Although the results should not be generalised to the entire UK market, this approach is consistent with established field-trial methodologies. Notably, despite variations in source and scope, the study's outcomes closely correspond with findings from other independent field trials.

Neither the Ofgem RHI dataset nor the published EoH datasets record compressor control type, and therefore performance cannot be disaggregated by the presence of speed-controlled (inverter-driven) compressors.

5. Discussion

5.1. Market context

The analysis of Ofgem data and the EoH Demonstration Project shows that heat pumps are capable of strong performance. The results show that 67% of all GSHPs and nearly 30% of all ASHPs operate at SPF 3.0 ($Ofgem_{H4}$) or higher, while one third of GSHPs and 8% of ASHPs achieve SPF 3.5 or above. These findings provide solid benchmarks for potential performance. The results demonstrate that the technologies are sound and scalable, but they also suggest that the challenge is consistency of performance across the market. As O'Hegarty et al. argue, the question is not whether we should install heat pumps "but how can we make sure to install heat pumps to maximise their potential" [11].

These findings are consistent with earlier UK and international studies that have identified a significant performance gap between in-situ and design SCOP performance [5,11]. The present analysis provides further robust evidence of the gap due to the sample sizes and the cross-validation between the two datasets obtained using different methodologies.

The performance gap directly impacts the economic value proposition. Recent analysis by the Resolution Foundation found that most families in the UK can expect to see their energy bills increase slightly when replacing a gas boiler with a heat pump, with energy costs increasing, on average, by £32 per year, and just one-third are set to save on running costs [14]. Other studies on the financial impact of heat pump installation report similar findings [15] including the Energy Saving Trust [16]. The Foundation notes that heating costs are particularly important for poorer families, accounting for 7% of disposable incomes, compared to 1.5% for the richest [14]. This economic reality makes accurate performance prediction critical for consumer protection and market confidence.

One weakness of the Resolution Foundation study is that the average financial impact assumes that all heat pumps perform with a coefficient of performance of 3.0 [14]. This is slightly higher than the average SPF found in both the analysis of Ofgem data and the EoH study and, as Section 3 sets out, many households experience SPFs much lower than 3.0.

Heat pump tariffs may reduce energy expenses for some consumers, but OVO Energy's withdrawal of its Heat Pump Plus tariff in February 2026 – discontinued with no replacement – has narrowed consumer choice and may indicate market instability [17]. According to the Resolution Foundation, time-of-use tariffs can result in savings, but their scalability may be limited as unit rates can be high at peak times, which will not suit all consumers. Heat Pump Plus was notable precisely because it avoided tariff complexity, offering a simple flat-rate rebate that reduced heat pump running costs by nearly half. Other options are only available for fixed 12-month terms such as Good Energy's Renewable Heat Pump Tariff [18]. The remaining options from Octopus, EDF, E.ON and British Gas require smart meters and active energy management to realise savings which are meaningful barriers for many households.

5.2. The Performance Gap: Causes and Implications

The performance gap has multiple potential causes. It is standard practice to compare design SCOP values with in-situ performance measured at the SPF_{H2} boundary, as both metrics use a comparable system boundary and represent heat-pump efficiency over an annual period. Comparing design SCOPs with the SPF_{H4} boundary is also valid since SPF_{H4} reflects the performance outcome experienced by consumers.

SCOP is a product metric derived under laboratory test conditions, while the SPF_{H2} boundary excludes several system components that are included within SPF_{H4}. Evidence from the EoH data (Table 2) indicates that the observed performance gap stems from factors well beyond these boundary-related differences. The EoH heat pump performance data shows a performance gap of 23% between the design SCOP and SPF_{H4} and a still-substantial gap of 19% between the design SCOP and SPF_{H2}.

Supporting evidence provided by the EoH team found system optimization improved some installations [19] using a range of strategies to address flow-temperature control, over-zoning, balancing, inappropriate set-points, DHW settings, and commissioning quality. All these optimisation activities improved outcomes without hardware changes, illustrating the potential of *measurement-led* quality assurance. The authors also said that poor installation quality also impacts energy consumption and efficiency. 20% of the installations in the EoH field trial were subject to quality assurance checks and inconsistencies in installation quality were highlighted. The authors note, however, that outside of the field trial context, consumers cannot always recognise poor installation practice.

O'Hegarty et al. identified five other possible explanations for the performance gap, and all relate to the European testing standards described in 2.2 [11]:

- **Temperature zone approximation.** The three standardised climate profiles used in air-to-water heat pump performance testing do not accurately reflect local conditions in regions such as the UK and Ireland, meaning that labelled SCOP values may misrepresent real-world efficiency in these climates.
- **Assumed relative humidity:** Standard test conditions fix relative humidity at specific setpoints, whereas actual humidity varies widely in operation; higher humidity increases defrosting frequency and thus reduces seasonal performance, contributing to lower in-situ efficiency than laboratory SCOP values suggest.
- **Assumed application temperature:** The SCOP calculation under EN 14825 is based on fixed indoor temperature assumptions that may differ substantially from real household settings.
- **Underestimated or neglected pipe heat losses:** Standard tests assume short, fully insulated pipe runs of 7.5 m or less, whereas real installations often have longer or poorly insulated connections.
- **Underestimated effects of cycling:** The simplified degradation coefficient (Cd) in EN 14825 does not fully capture the efficiency losses from frequent on/off cycling so field-monitored data typically reveal lower seasonal performance than the standard predicts.

Research also identifies mistakes in heat pump sizing as a likely factor in poor performance. An investigation carried out by Elmhurst Energy, Build Test Solutions and Veritherm in 2023 reported that the average difference between the *calculated* and *measured* heat loss in UK dwellings was 35% with a significant bias towards overestimating heat loss [20][21]. The study found that heat loss calculated using the UK's established methodology (BS-EN 12831) used for heat pump sizing overestimated that heat loss for 59% of houses. As oversized heat pumps have lower running efficiency, the authors argue that heat pump sizing should be based primarily on heat loss *measurement* and that UK standards should be updated.

Consumer behaviour can affect efficiency, but factors like high comfort setpoints are hard to disentangle from technical issues discussed above. Consumers may, for example, use sub-optimal control strategies to correct problems beyond their ability to resolve. The EoH Optimisation Report recommends using *system monitoring* for optimisation strategies to achieve consumer comfort goals while also maximising efficiency [19].

6. Recommendations and conclusion

6.1. The SCOP metric is not suitable for system performance forecasting

Consumer confidence in heat pumps depends on the ability to make informed decisions based on realistic, context-specific performance information. The evidence suggests, however, that there is a serious disconnect between design SCOPs and consumer experience. The rules that underpin the UK's Boiler Upgrade Scheme (BUS) incentive, for example, insist that all heat pumps must have a minimum SCOP of 2.8 and the government has said that the "system efficiency will be evidenced through the MCS certificate submitted by the installer [22]." However, over 50% of the ASHP installations in the EoH study performed with an SPF_{H4} below that 2.8 benchmark and 37% were below that SPF_{H2} 2.8 benchmark. As noted in this paper, compliance with minimum SCOP requirements does not ensure SPF performance. This demonstrates that the minimum efficiency standards of the BUS scheme have not prevented low and very low performance.

Given the performance issues identified in this paper, the UK's compliance-based regulatory model should evolve into an evidence-based, adaptive, and outcome-oriented framework. This new approach should be explicitly grounded in the principles of regulatory science – emphasizing empirical evidence, continuous monitoring, and structured feedback loops to guide policy decisions and ensure that interventions deliver measurable and verifiable real-world outcomes.

The MCS no longer uses SCOP as the basis for its performance estimation methodology [9]. Nevertheless, as outlined in Section 2, discourse regarding performance remains centered on the SCOP metric. This emphasis is partly attributable to the mandatory inclusion of these figures on energy labels and the alignment of company marketing with those labels. This emphasis on SCOPs as a proxy for performance is unsustainable and, whilst identifying solutions to the causes of performance variation discussed in Section 5 is beyond the scope of this paper, the need for improved performance monitoring of installations is self-evident.

6.2. Minimum performance guarantees should be used to improve consumer confidence

Carmichael argues that the UK's heat-pump strategy must evolve from modelled to measured performance, noting that reliance on SCOP-based design forecasts has limited accountability and can undermine consumer confidence. He recommends standardised, low-cost in-situ monitoring for all new installations, with verified results feeding into a national performance database to inform regulation, training, and consumer guidance. Such measurement-based assurance, he contends, would enable the introduction of minimum verified performance thresholds for publicly funded installations and align incentives with real-world outcomes [23].

Carmichael's proposals are echoed in key responses to the EoH Demonstration Project findings. The managing contractor of the EoH Demonstration Project, Energy Systems Catapult, has similarly advocated outcome-based assurance, stating that "incentivising contractual arrangements such as performance guarantees and assigning outcome responsibility to contractors will help to avoid contractors following a 'fit-and-forget' approach and will ensure a better outcome for consumers" [24]. The *EoH Optimisation Report* also recommended new business models and redress mechanisms to promote installation quality, and it notes that optimization strategies depend on on-going system performance monitoring [19].

In parallel, the Resolution Foundation proposed a *minimum system COP of 3.25*, estimating this would ensure lower running costs than a gas boiler for around 65 % of households [14]. Collectively, these proposals highlight the need for outcome-based regulation that prioritises consumer protection and accountability to build consumer trust.

Regulators should not only ensure that installers comply with the new performance estimation methodology deployed by MCS, but they should also consider how to implement systematic national monitoring. This could be implemented through an anonymised case-study database encompassing diverse housing archetypes, allowing stakeholders to learn from verified outcomes and identify performance drivers [22]. Regulators and other stakeholders could use the database to improve forecasting methodologies to incorporate building-specific factors and to encourage installer or third-party optimisation in partnership with consumers. Regulators should use incentive mechanisms for ongoing system monitoring and data sharing.

In response to concerns about installation quality and performance shortfalls, some UK installers have already adopted an outcome-focused approach using minimum performance guarantees. Heat Geek, for example, provides a five-year assurance that each system will perform at least as efficiently as quoted [25].

The analysis presented here demonstrates that while heat-pump technology is fundamentally sound and capable of high efficiencies, a persistent performance deficit threatens to undermine public confidence and slow the decarbonisation of heat. Regulatory and policy frameworks should therefore evolve to prioritise measured performance, outcome accountability, and consumer protection through verified efficiency standards, post-installation monitoring, and performance-based incentives. Policies that *de-risk installations for consumers* – by ensuring transparent information, enforceable quality standards, and verified in-situ outcomes – will be crucial to restoring trust and making heat pumps a competitive alternative to gas boilers.

Acknowledgements

Research funded by UK's DESNZ, Renewable Energy Consumer Code (RECC), and the Ground Source Heat Pump Association (GSHPA). Authors also thank DESNZ for providing the dataset. Research conducted with support from Excel Specialist Jamie Newman. Authors acknowledge valuable insights from Electrification of Heat Demonstration Project and Energy Systems Catapult.

References

[1] Committee on Climate Change (CCC). (2019). Net Zero: The UK's contribution to stopping global warming – Technical Report. London: CCC, p. 84.
[2] Rowe, G. and Rankl, F. (2024). Housing and net zero: Research Briefing (CBP-8830). House of Commons Library, 8 July 2024. Available at: https://researchbriefings.files.parliament.uk/documents/CBP-8830/CBP-8830.pdf
[3] Committee on Climate Change (2024). Progress in reducing emissions: 2024 Report to Parliament. July 2024. ISBN 978-1-5286-4991-9

[4] UK Government. (2023, 21 November). Heat pumps for domestic heating: What impact can heat pumps have in domestic heating today and how might that change over time as technology improves? London: Department for Energy Security and Net Zero. Available at: https://www.gov.uk/government/publications/heat-pumps-for-domestic-heating/what-impact-can-heat-pumps-have-in-domestic-heating-today-and-how-might-that-change-over-time-as-technology-improves-html
[5] Lowe, R., Summerfield, A., and Oreszczyn, T. (2017). Report on analysis of heat pump data from the Renewable Heat Premium Payment (RHPP) scheme. London: UCL Energy Institute for the Department for Business, Energy & Industrial Strategy (BEIS). Issued March 2017. Available at: https://www.gov.uk/government/publications/detailed-analysis-of-data-from-heat-pumps-installed-via-the-renewable-heat-premium-payment-scheme
[6] Energy Systems Catapult Ltd. (2024). Electrification of Heat Demonstration Project: Heat Pump Performance Data Analysis Report. Bimingham: Energy Systems Catapult. Available at: https://es.catapult.org.uk/report/electrification-of-heat-summary-reports-and-datasets/
[7] MCS. (2017). MIS 3005 Requirements for mcs contractors undertaking the supply, design, installation, set to work, commissioning and handover of microgeneration heat pump systems. Issue 5. April 2017.
[8] Meek, Colin; Hill, Frances (2022). Heat pumps and the UK's heat decarbonisation: Lessons from an Ofgem dataset of more than 2,000 domestic installations. Leeds Beckett University. Conference contribution. https://hdl.handle.net/10779/leedsbeckett.28946057.v1
[9] MCS. (2025). Heat Pump System Performance Estimate Standard and Tool (MCS 031). Information and Guidance for Installers.
[10] Project Ambient. (2025). Ambient website and calculator. Project Ambient CIC. Retrieved 10.11.2025, from https://www.projectambient.org/ambient-cost-calculator
[11] O'Hegarty R., Kinnane O., Lennon D., Colclough S. (2022). Air-to-water heat pumps: Review and analysis of the performance gap between in-use and product rated performance. Renewable and Sustainable Energy Reviews. Volume 155, 2022, 111887, ISSN 1364-0321. https://doi.org/10.1016/j.rser.2021.111887
[12] Competition and Markets Authority. (2023). Consumer protection in the green heating and insulation sector – Final Report (31 May 2023). London: CMA.
[13] Ofgem. (2018). RHI Metering Guide V3.0. London: Ofgem.
[14] Leather, Z. & Marshall, J. (2025, April). Turning up the heat: Making the home heating transition work for low-income households. London: Resolution Foundation. Available at: https://www.resolutionfoundation.org/publications/turning-up-the-heat-making-the-home-heating-transition-work-for-low-income-households/
[15] Harrington, N., Goodlad, R., Webb, J., Wade, F., & McKenna, E. (2024). *Motivations and Barriers Associated with Heat Pumps.* Glasgow: UK Collaborative Centre for Housing Evidence (CaCHE), University of Glasgow. Available at: https://housingevidence.ac.uk/publications/motivations-and-barriers-associated-with-heat-pumps
[16] Energy Saving Trust. (2025, last updated 1 October 2025). *Heat pumps: how they work, costs and savings*. [Web page]. Retrieved 10.11.2025 from https://energysavingtrust.org.uk/advice/in-depth-guide-to-heat-pumps/
[17] OVO Energy (2026) Heat Pump Plus: An important update. Retrieved March 2026, from https://www.ovoenergy.com/heat-pump-plus
[18] Good Energy. (n.d.). Good Energy Heat Pump. Tariff Rates. Retrieved March 2026, from https://www.goodenergy.co.uk/
[19] Energy Systems Catapult Ltd. (2024). Electrification of Heat Demonstration Project: Heat Pump Optimisation Report. Birmingham: Energy Systems Catapult. Available at: https://es.catapult.org.uk/report/electrification-of-heat-summary-reports-and-datasets/
[20] Build Test Solutions. (2023). MEASURED: Improving the survey and design process – Final report. London: Build Test Solutions. Available at: https://assets.publishing.service.gov.uk/media/68834b68c02b468fa1ae092c/Build_Test_Solutions.pdf
[21] Elmhurst Energy. (2023, December 19). Accurate heat loss measurement is vital to heat pump sizing [Blog post]. Available at: https://www.elmhurstenergy.co.uk/blog/2023/12/19/accurate-heat-loss-measurement-is-vital-to-heat-pump-sizing/
[22] UK Government. (2021, October). *Future Support for Low Carbon Heat: Government response to the Clean Heat Grant proposals within 'Future support for low carbon heat' consultation* (Boiler Upgrade Scheme). London: HM Government. Available at: https://assets.publishing.service.gov.uk/media/616d8ac18fa8f52986e61ced/clean-heat-grant-government-response.pdf
[23] Carmichael, R., (2022) Accelerating the transition to heat pumps: measuring real-world performance and enabling peer-to-peer learning. An Energy Futures Lab Briefing Paper. Imperial College London. Available at: https://www.imperial.ac.uk/energy-futures-lab/ reports/briefing-papers/paper-10/
[24] Energy Systems Catapult Ltd. (2024). Electrification of Heat Demonstration Project Summary Report. Birmingham: Energy Systems Catapult. Available at: https://es.catapult.org.uk/report/electrification-of-heat-summary-reports-and-datasets/
[25] Heat Geek Installations Ltd. (n.d.). Installation terms: Homeowner & Guarantee. Retrieved 10.11.2025, from https://www.heatgeek.com/tncs/homeowner-agreement