UCL **ENERGY** INSTITUTE

FINAL REPORT ON ANALYSIS OF HEAT PUMP DATA FROM THE RENEWABLE HEAT PREMIUM PAYMENT (RHPP) SCHEME

Issued: March 2017

RAPID-HPC Authors: Robert Lowe, Alex Summerfield, Eleni Oikonomou, Jenny Love, Phillip Biddulph, Colin Gleeson, Lai-Fong Chiu and Jez Wingfield

Table of contents

List of figures

List of tables

Nomenclature

PERFORMANCE EFFICIENCY NOMENCLATURE

COP   Heat pump (HP) coefficient of performance

SPF_{Hn} HP seasonal performance factor for heating at SEPEMO boundary Hn

MONITORED VARIABLES

Eb	Electricity for whole system boost only
Edhw	Electricity for domestic hot water (typically an immersion heater)
Ehp	Electricity for the heat pump unit (may include a booster heater and circulation pump)
Esp	Electricity for boost to space heating only
Fhp	Flow rate of water from heat pump (may be space heating only)
Fhw	Flow rate of water to DHW cylinder (if separately monitored)
Hhp	Heat from heat pump (may be space heating only)
Hhw	Heat to DHW cylinder (if separately monitored)
Tco	Temperature of water leaving the condenser
Tin	For ASHP: Temperature of refrigerant leaving the evaporator
	For GSHP: Temperature of ground loop water into the heat pump
Tsf	Flow temperature of water to space heating
Twf	Flow temperature of water to cylinder

(Note that external temperature, Tex, was not measured directly. Data from a publicly available database were used in the analysis.)

RHPP ENERGY AND POWER UNITS

Energy	J	Joule	SI unit of energy
Energy	kWh	3.6 MJ	Customary unit of energy for residential energy use
Energy	MWh, GWh	3.6 GJ, 3.6 TJ	
Power	W	Watt, J/s	SI unit of power and heat flow
Power	Wh/2 minutes	30 W	Base unit of energy for monitored data in RHPP trial, limit of resolution of power – note that power and heat have been recorded at 2 minute intervals
Power	kWh/year	3.6 MJ/year 0.11416 W	Customary unit for rate of residential energy use
Power	kW	1000 W	Typical unit for measurement of heating system ratings

KEY ACRONYMS AND ABBREVIATIONS

DECC	Department of Energy and Climate Change (became part of the Department of Business, Energy & Industrial Strategy on 14th July 2016)
EST	Energy Saving Trust
Preliminary Assessment	Preliminary assessment of the RHPP data performed by DECC (Wickins, 2014)
RAPID-HPC	Research and Analysis on Performance and Installation Data – Heat Pump Consortium
RHPP	Renewable Heat Premium Payment Scheme
MCS	Microgeneration Certification Scheme - a nationally recognised quality assurance scheme, supported by the DECC. MCS certifies microgeneration technologies used to produce electricity and heat from renewable sources.
MIS	Microgeneration installation standards. MIS 3005 set out requirements for MCS contractors undertaking the supply, design, installation, set to work, commissioning and handover of microgeneration heat pump systems.
SEPEMO	SEasonal PErformance factor and MOnitoring

Context

The RHPP policy provided subsidies for private householders, Registered Social Landlords and communities to install renewable heat measures in residential properties. Eligible measures included air and ground-source heat pumps, biomass boilers and solar thermal panels.

Around 14,000 heat pumps were installed via this scheme. DECC[1] funded a detailed monitoring campaign, which covered 700 heat pumps (around 5% of the total). The aim of this monitoring campaign was to provide data to enable an assessment of the efficiencies of the heat pumps and to gain greater insight into their performance. The RHPP scheme was administered by the Energy Savings Trust (EST) who engaged the Buildings Research Establishment (BRE) to run the meter installation and data collection phases of the monitoring program. They collected data from 31st October 2013 to 31st March 2015.

RHPP heat pumps were installed between 2009 and 2014. Since the start of the RHPP Scheme, the installation requirements set by MCS standards and processes have been updated.

DECC contracted RAPID-HPC to analyse this data. The data provided to RAPID-HPC included physical monitoring data, and metadata describing the features of the heat pump installations and the dwellings in which they were installed.

The work of RAPID-HPC consisted of cleaning the data, selection of sites and data for analysis, analysis, and the development of conclusions and interpretations. The monitoring data and contextual information provided to RAPID-HPC are imperfect and the analyses presented in this report should be considered with this in mind. Discussion of the data limitations is provided in the reports and is essential to the conclusions and interpretations presented. This report does not assess the degree to which the heat pumps assessed are representative of the general sample of domestic heat pumps in the UK. Therefore these results should not be assumed to be representative of any sample of heat pumps other than that described.

Acknowledgements

---

[1] The Department of Energy and Climate Change (DECC) merged with the Department for Business, Innovation and Skills (BIS) in July 2016, to create the new Department for Business, Energy & Industrial Strategy (BEIS)

1 Introduction

1.1 RHPP Field Trial

The opportunity to conduct a large-scale field trial of heat pumps in the UK arose from the Renewable Heat Premium Payment (RHPP) policy that provided subsidies for private householders, Registered Social Landlords, and communities to install renewable heat systems in residential properties. Eligible technologies included air and ground-source heat pumps, biomass boilers, and solar thermal panels. The scheme was administered by the Energy Saving Trust (EST) and operated from August 2011 to March 2014. The RHPP policy provided subsidies to around 14,000 heat pumps, including some heat pumps installed prior to the start of the scheme (and as early as 2009).

The Department of Energy and Climate Change (DECC), which has since been incorporated into the new Department of Business, Energy and Industrial Strategy (BEIS), funded a detailed monitoring campaign with the aim of assessing the performance of the heat pumps. DECC contracted The Energy Savings Trust (EST), which contracted the Buildings Research Establishment (BRE) to run the meter installation and data collection phases of the monitoring programme. The resultant RHPP Field Trial covered a total of 700 heat pump sites, with 2 minute heat and electricity data collected from 31st October 2013 to 31st March 2015.

DECC contracted the RAPID-HPC to analyse this data. Work, spanning just over two years, began in December 2014 and finished in March 2017. The primary outputs from this work are five reports:

- the _RHPP Performance Variations Report[2]_, which applies primarily statistical analysis to the physical monitoring data to provide quantitative estimates of SPF and associated uncertainty, and presents exploratory analysis of reasons for observed variations in performance;

- the _RHPP MCS Compliance Report_, which analyses the extent and effect of compliance with MIS 3005 as it applied to the installation of heat pumps in the RHPP Field Trial;

- the _RHPP Case Study Report_, which describes detailed studies undertaken on 21 individual sites within the RHPP Field Trial sample;

- the _RHPP Bias Errors report_, which assesses measurement errors and their impacts on estimated efficiencies;

---

[2] These are abbreviated titles. The full titles of the first three reports are given in the References section of the present report.

- and finally, this, the <u>*RHPP Final* Report</u>, which attempts to bring together insights from all three of the above reports in a single short document.

1.2   Seasonal Performance Factor

Seasonal Performance Factor, or SPF, is an assessment of the efficiency of a heat pump system, i.e. heat out divided by electricity in.

HP system boundaries are fundamental to the evaluation of annual HP performance using monitored data, with specific parameters SPFH1, SPFH2, SPFH3 and SPFH4 explained and applied in previous work reporting on the SEPEMO project (Riviere et al., 2011) and the second phase of EST's HP field trial (Dunbabin et al., 2013). The significance of using different system boundaries has been discussed in previous literature, for instance on a field trial of HP systems in Germany analysed by a team at the Fraunhofer ISE Institute (Miara et al., 2011), and by Gleeson and Lowe (2013).

The SEPEMO methodology (Appendix A1) starts from the core HP components only[3] (SPFH1), with expanding boundaries covering the supply air fan or ground loop pump power into the HP (SPFH2), backup heaters, including electric immersion for domestic hot water if present (SPFH3) and finally, system circulators or pumps (SPFH4). For the purposes of comparison with the first phase of the EST trials (Dunbabin *et. al* 2012, 2013) an "H5" boundary has been added, which accounts for the heat losses from the hot water cylinder and allows heat output to be measured at the outlet of an integrated hot water store. Note the relationship between the higher index number and lower numerical value for SPF for the same installation.

The clarity of the SEPEMO method and its applicability to estimating the renewable heat from HP systems of various types, as well as the need for standard approaches for performance indicators has resulted in its adoption across Europe. Notably, for the purposes of estimating the amount of renewable energy generated by heat pumps for the EU Renewable Energy Directive, the EU Commission Decision of 1st March 2013 defines the relevant measure of efficiency to be SPFH2, i.e. the efficiency of the heat pump without the inclusion of electricity used for backup or distribution circulation pumps. The Directive considers heat pumps as providing renewable energy provided that SPFH2 $\geq$ 2.5[4].

---

[3] Compressor, condenser, expansion valve, evaporator.

[4] The criterion for heat pumps to be considered "renewable" under the EU Commission Decision of 1st March 2013 is based on the primary energy efficiency of the EU electricity system. Currently, the assumed efficiency is 40%, but this is expected to increase significantly over time as the proportion of primary electricity increases. The effect of this, if followed through, would be that the limiting SPF would reduce.

2 Methods and Limitations

2.1 Field trials: practicalities and expectations

The key objective of a field trial is to gain understanding of HP system operation and performance *in-situ*, which augments findings from the controlled and short-term testing possible in more closely controlled factory and laboratory settings. The objective not only concerns the overall typical estimated performance of the sample, but also to identify the spectrum of issues, technical and socio-technical, some of which will be wholly unexpected, that are only likely to emerge when a large number of HP systems are installed and monitored in diverse, real-world settings.

The key metrics used for the study are Seasonal Performance Factors (SPFs), which by definition measure the energy performance at various system boundaries *over a continuous 12 month period.* SPF is an extension of the concept of coefficient of performance (CoP), which is measured over shorter operational periods. COP would be the metric that would typically by used to measure performance in the laboratory or as the basis for investigations of variability in performance over timescale varying from diurnal to seasonal. Evaluation and interpretation of SPFs in this project has taken place against the following background:

- The installation arrangement of physical monitoring systems in the monitored HP systems is highly variable. The layout of the heating system and corresponding sensors for each site is documented through a 'schematic', which may itself be subject to error. There is limited redundancy in sensors in the RHPP project[5], which means that there is limited scope to use the monitored data the detect and fix monitoring system problems. Indeed such issues may only become apparent relatively late in the data analysis phase.

- Unlike an intervention study, where a problematic system would be repeatedly examined and adjusted until the issue was identified and resolved, the HP systems in the field trial were not subject to quality control procedures that were above and beyond those normally part of the installers' and the manufacturers' protocols.

The distinct features of this field trial study design has three immediate consequences. The typical SPF performance of HP systems should not be expected to be the same as those obtained under controlled conditions, where there is scope to address metering and operational issues as they arise. Second, due to the wider range of factors potentially at work, there will also be greater variation seen in the distribution SPF performance of HP systems across the sample of sites. Last, the data used for the analysis of each

---

[5] There is a trade-off between redundancy, complexity and cost (including cost of quality assurance) in any physical monitoring project.

site will contain some 'anomalies', such as missing data, which may be due to metering errors or real operational issues. They may also be due to legitimate reasons that would not occur in other settings, such as the occupants deciding to turn the HP system off for a spell. To automatically disqualify or omit any site from further analysis, simply on the basis of detecting a transient data anomaly, would defeat the purpose of a field trial of in-situ HP system operation. While there is always scope for debate around the methods of detection of data errors and the extent or frequency of an issue, such as missing data across the year, needed before a site is omitted from contributing to the SPF evaluation, there should be no expectation that the monitoring data used in the analysis must be perfect. This is all the more relevant given the need for a sample size large enough to support statistical inference with respect to relative performance of different groups of HP systems and impact of site characteristics.

2.2   Complementing the statistical analysis

The work of RAPID-HPC has not relied solely on statistical analysis of remotely monitored data. The project included two complementary strands of work:

- an analysis of compliance with the MCS; this included detailed analysis of metadata and associated procedures.

- a case study, involving 21 cases selected from the *preliminary and unpublished dataset*.

These two additional strands of work shed significant additional light on the quality of the metadata and the technical context within field trial dwellings.

In particular, the case study involved site visits, typically of 2-3 hours' duration, which allowed semi-structured interviews with occupants and a technical survey of the case study dwellings and their heating systems. The case studies provided rich additional insight into dwellings, heating systems, control systems, monitoring systems, events that befell some of the systems during the monitoring period, and finally, occupants' lifestyles, heating control strategies and satisfaction with their systems. All of this significantly enriched the overall interpretation of data and results from statistical analysis.

3   Data collected, data quality and filters applied

This section describes the data collected, the data that were not collected and the filters that have been applied to the data to define sub-samples for statistical analysis.

3.1   Monitoring data collected

Table 3-1 shows the complete set of parameters in the monitored data used to calculate the SPFs, though it should be noted that different sites had different combinations of parameters according to the schematic (or monitoring layout) that was applicable to that installation and plumbing arrangement.

Table 3-1. The complete set of parameters included in the monitored data

Parameter	Description
Eb	Electricity meter for whole system boost only
Edhw	Electricity meter for domestic hot water (typically an immersion heater)
Ehp	Electricity meter for the HP unit (may include a booster heater and circulation pump)
Esp	Electricity meter for boost to space heating only
Fhp	Flow rate of water from HP (may be space heating only)
Fhw	Flow rate of water to DHW cylinder
Hhp	Heat meter from HP (may be space heating only)
Hhw	Heat meter to DHW cylinder
Tco	Temperature of refrigerant leaving the condenser
Tin	For ASHPs: Temperature of refrigerant leaving the evaporator For GSHPs: Temperature of ground loop water into the HP
Tsf	Flow temperature of water to space heating
Twf	Flow temperature of water to cylinder

Figure A 2 and Figure A 3 in Appendix A2 provide examples of two simple schematic diagrams (for an air source and ground source HP respectively) that illustrate the location of monitoring points corresponding to the monitored parameters in any given heat pump system. Full details of the monitoring programme, including the overall monitoring philosophy and considerations of sensor resolution, can be found in the Preliminary Assessment report (Wickins, 2014).

Data in the metadata file include:

a) Heat pump net capacity

b) MCS certificate number

c) Installer's estimates of annual heat demand

d) Heat pump type

e) Tenure

f) Previous heating fuel (for retrofitted owner-occupied properties only).

In a few cases, additional data were available, e.g. circulation pump power and settings, data on supplementary heaters such as log fires etc.

Data that were not collected include:

a) Temperatures within the dwellings. (UK Met data have been used to estimate external temperatures.)

b) Information on ground loop design for GSHPs.

c) Information on radiator sizes (except for the case studies).

d) Information on the design of underfloor heating systems.

e) Size of domestic hot water tanks (although this was recorded in some of the case studies).

f) Total delivered energy to the dwelling (electricity and, where applicable, mains gas measured by utility meters, and records of deliveries of coal, oil and LPG)[6].

---

[6] Though not all of these will be relevant in every case, data on total delivered energy in principle allows multiple independent checks on the integrity of monitoring systems, use of secondary heating etc. Use of smart meter data is likely to increase the range and analytical value of such checks.

3.2   The B2 dataset

Once it was confirmed which sites had the correct data-streams recorded, a number of simple filters for data quality and completeness were applied for the selection of data to be used in the analysis of SPFs. Those sites included in analysis were required to have 13 consecutive months of data, with at least five days of monitoring from each month. The 13 month period with the most stable water flow rate between the first and last month was selected for each site. This resulted in the selection of Sample B2 with 318 ASHPs and 99 GSHPs. This dataset showed SPFs between 0 and 5.5.

3.3   Data quality issues and the cropped B2 dataset

RAPID-HPC then filtered out data for which 1.5<SPFH4<4.5 on the grounds that these data are likely to have been influenced by monitoring anomalies. This procedure had previously been adopted by Pedersen and Jacobsen, 2011. It was considered the simplest and most transparent approach, and preferable to developing and testing numerous bespoke algorithms to detect each type of potential monitoring anomaly.

The dataset produced by this succession of filters is referred to as the "Cropped B2 dataset". Its characteristics shown in Table 4-2 below:

Table 3-2. Characteristics of the Cropped B2 dataset

SPF	Heating service	HP type	N	Mean (95% CI)	Median (IQR)	% with SPFH2>=2.5 (95% CI)*
SPFH2	Overall	ASHP	292	2.64 (2.60, 2.70)	2.65 (2.33-2.95)	62% *(56%-68%)*
		GSHP	92	2.93 (2.80, 3.06)	2.81 (2.63-3.14)	80% *(72%-88%)*
SPFH2	SH	ASHP	292	2.72 (2.66, 2.78)	2.74 (2.36-3.09)	**
		GSHP	92	3.03 (2.86, 3.2)	2.89 (2.59-3.34)	**
SPFH2	DHW	ASHP	284	2.3 (2.24, 2.36)	2.31 (2.05-2.56)	**
		GSHP	78	2.7 (2.56, 2.85)	2.71 (2.29-2.99)	**
SPFH4	Overall	ASHP	293	2.41 (2.37, 2.46)	2.44 (2.15-2.67)	**
		GSHP	92	2.77 (2.66, 2.89)	2.71 (2.48-3.02)	**

* 95% CI were calculated as p ± 1.96 * √(p*(1-p)/n), where p is the proportion identified as ≥2.5.

** = Not applicable.
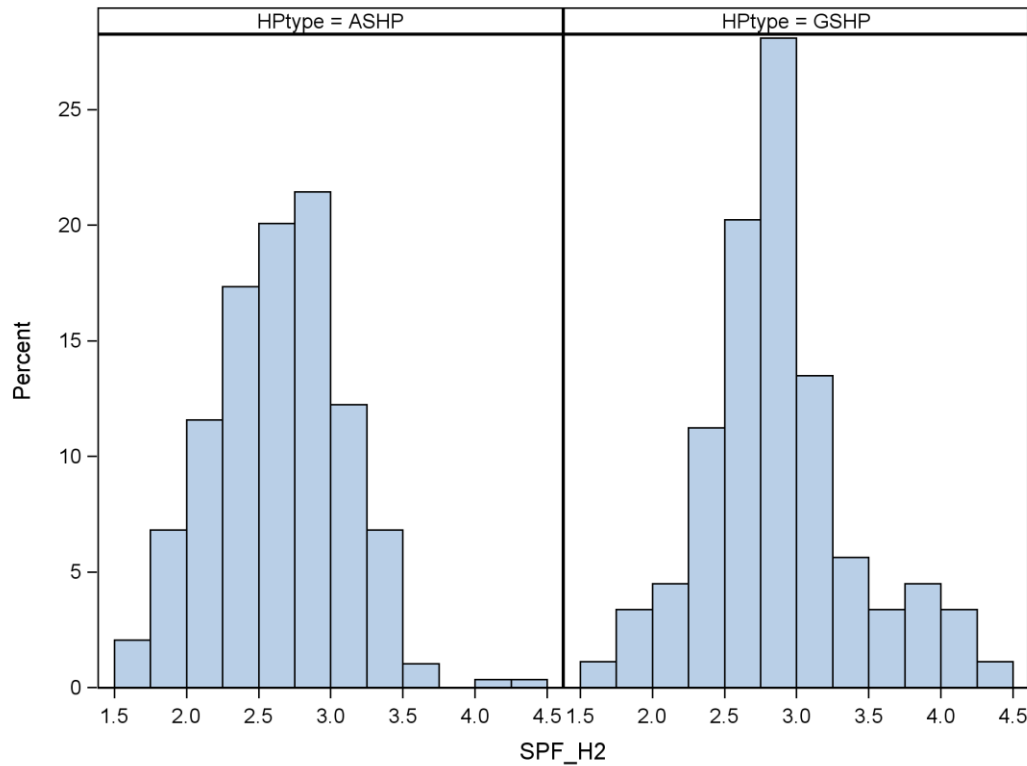
Figure 3-1 Histogram of SPFH2 values for the cropped B2 dataset – left ASHP, right GSHP

For sites within this filtered dataset, detailed assessments of the influence of metering errors on estimates of seasonal performance factor were carried out. Some types of error have the effect of increasing SPF, while others have the effect of decreasing it. Where possible, these errors and their prevalence were quantified; most only affect a sub-set of sites. It is estimated that missing heat meter data is likely to have had an effect of <4% on the median SPFH2 across the sample, although estimates of SPFH2 for individual sites may be more affected. Many sites will be affected by a calibration error, because the Sontex heat meters used in the project were calibrated for water, rather than glycol[7]. The analysis indicates that as a result, SPFs will be over-estimated by 4-7%. This correction factor has not been applied to the estimated SPF's in RAPID-HPC reports. The detail of the calculation is set out in RAPID-HPC (2017e). Appendix A3 provides further detail on monitoring system errors.

3.4 Alternative methods of filtering data

Instead of filtering out sites for which 1.5<SPFH4<4.5, alternative filters could have been used. Table 3-3 presents the effect of alternative filters on the estimated median SPF's and the percentage meeting the criterion for being considered "renewable". These alternative filters are:

1. No filter (analysis takes place on the whole of sample B2).

2. Cropped B2: exclude sites from B2 sample where SPFH4 was < 1.5 and > 4.5

3. B2 Tukey: exclude sites from B2 sample according to the standard Tukey definition of outliers:

---

[7] It is possible that primary circulation circuits at some sites may not have been treated with antifreeze.

P25 - 1.5*IQR and P75 + 1.5*IQR, where P25 and P75 refer to the 25th and 75th percentile respectively and the Inter-Quartile Range is defined as IQR = P75 – P25. For ASHPs these were calculated for SPFH4 as < 1.19 and > 3.52, and for GSHPs as <1.47 and > 3.98 (Tukey, 1977).

It should be noted that while, for ASHPs, the B2 Tukey sample is slightly less restrictive than the cropped B2, for GSHPs the two samples are almost identical. A number of other statistical methods are available to define outliers, such as the Median Absolute Deviation (MAD). However these all tend to lie outside of the Tukey boundaries, so their impact will be such that the results will lie between the B2 sample and the B2 Tukey sample. Also note that more aggressive filters will further reduce the sample size and so increase the statistical uncertainty around estimates of heat pump performance based on the remaining sites.

Table 3-3. Comparison of SPFs obtained for different filters applied to Sample B2

HP type	Sample	N	SPFH2 Median (IQR)	Percentage SPFH2≥2.5 (95% CI)*
ASHP	B2	318	2.63 (2.24, 2.94)	*58% (52-64%)*
	cropped B2	292	2.65 (2.33, 2.95)	*62% (56%-68%)*
	B2 Tukey	305	2.63 (2.28, 2.94)	*60% (54%-66%)*
GSHP	B2	99	2.81 (2.52-3.16)	*77% (68-86%)*
	cropped B2	92	2.81 (2.63-3.14)	*80% (72-88%)*
	B2 Tukey	90	2.81 (2.53-3.09)	*79% (71-87%)*

* 95% CI were calculated as $p \pm 1.96 * \sqrt{(p*(1-p)/n)}$, where p is the proportion identified as ≥2.5, where these estimates for the size of uncertainty are less appropriate for proportions of more than ~80%.

This analysis demonstrates that the median SPFH2 values are barely affected by the alternative filters used and can therefore be considered robust for this sample.

The estimates of the proportion of ASHP and GSHP that meet the Renewable Energy Directive definition of "renewable" are affected by the choice of filter, but, crucially, the confidence intervals for this statistic overlap. The results here indicate that for the range of filters so far tested, around two thirds of ASHPs and four out of five of GSHPs had sufficiently high SPFs to be classified as renewable sources of energy, under the terms of the Renewable Energy Directive. However, the robustness of this conclusion to variations in method renders it essentially qualitative.

This suggests the potential for policies and/or business models that aim to identify and improve the one in three ASHPs and one in five GSHPs that, if this sample is indicative of heat pumps more generally, do not appear to attain the level of performance defined by the Renewable Energy Directive.

4   Principal Results

We conclude from the above analysis that the principal results presented in this report for performance of HPs in the RHPP sample are not greatly affected by the choice of filter for removing outliers.

Using the cropped B2 dataset, we present below the minimum SPFs required to ensure that heat pumps make CO_2 savings relative to alternative fuels, and of the proportion of sites making CO_2 savings and bill savings relative to alternative fuels.

4.1   Minimum SPFH4's required to make CO_2 and bill savings relative to alternative fuels

Table 4-1 shows the SPFs that must be exceeded for the systems to make a) CO_2 savings and b) energy bill savings, with corresponding estimates for CO2 and bill savings for detached, semi-detached, and terrace dwellings which represent the majority of residential types in the UK. Bill savings, particularly, are very sensitive to tariffs; in each case the assumptions used are described in the table notes.

One key point is that for the purposes of calculation, the heat demand (based on median gas consumption in the three dwelling types) is assumed to remain unchanged with each alternative fuel. Thus it does not account for the potential of occupants to alter their heart demand, for instance in response to changes in energy costs.

Note that although terms such as 'take back' are often used to describe the combination of higher internal temperatures and higher demand for heat that result from the installation of a new technology, the use of such terms in the context of heat pumps may be problematic. This is because of the potential for complex interactions between intermittent/continuous heating, mean flow and return temperatures in heat distribution systems, and SPF. For heat pumps with good part load performance, it is possible in principle to shift from intermittent to continuous heating and simultaneously increase annual heat demand, and reduce electricity consumption and CO_2 emissions.

The efficiency requirements to ensure CO_2 savings are far less onerous than those to ensure cost savings. Since for all the fuels the SPFH4 needed to ensure CO_2 savings is below the minimum cut-point for the cropped Sample B2, 100% of the sites reach this standard.

Table 4-1. Fuel costs and efficiencies used to determine the SPF$_{H4}$ of an HP system (on standard tariff electricity) needed in order to outperform other fuels

Fuel[a]	Carbon intensity (gCO$_2$eq/kWh)	Fuel cost (p/kWh)	System efficiency	SPFH4 needed for reduction in CO$_2$ emissions	SPFH4 needed for fuel cost reduction
Electricity (standard)	324 [b,c]	13.86[d]	100%	1.00	1.00
Electricity (economy 7)	324 [b,c]	7.21[d]	100%	1.00	1.92
Coal	366 [b]	3.94[d]	60%	0.53	2.11
Oil	268 [b]	3.58[d]	84%	1.02	3.25
LPG	215 [b]	6.66[d]	85%	1.30	1.77
Gas	184 [b]	4.18[d]	85%	1.5	2.82

[a] Selected fuels potentially displaced by HP systems (using electricity on the standard tariff). The RHPP was not intended to fund the replacement of gas-fired heating with heat pumps, but gas is included in this and subsequent tables as the current best performer among non-renewable fuels.

[b] Source: Green Book supplementary guidance: valuation of energy use and greenhouse gas emissions for appraisal, DECC, last updated 11th December 2015, https://www.gov.uk/government/publications/valuation-of-energy-use-and-greenhouse-gas-emissions-for-appraisal See Data_tables_1-20_supporting_the_toolkit_and_the_guidance.xlsx Table 1.

[c] Note this is a long-tun marginal value, not a current grid average, which is significantly higher. - see DECC 2015b. (See accompanying tables Data_tables_1-20_supporting_the_toolkit_and_the_guidance.xlsx). This document suggests a continuous decline in carbon intensity for grid electricity over the coming decades, and a convergence of marginal and average values. The marginal intensity is projected to fall to around 0.282 in 2020 and to around 0.129 g(CO$_2$)/kWh in 2030.

[d] Source: http://www.energysavingtrust.org.uk/about-us/our-calculations (Fuel prices and carbon intensity, updated March 2016). Fuel prices could also be taken from reference (b) above; however, this does not include LPG or Economy 7 Tariff for electricity and so a single price source has been preferred.

Table 4-2. Median annual CO_2 savings (tonnes per dwelling) and percentage of HP systems expected to yield energy CO_2 savings relative to other fuels based on cropped Sample B2.

	ASHP	ASHP median SPFH4=2.44			GSHP	GSHP median SPFH4=2.71		
Displaced fuel[a]	%[b]	Detached	Semi-det.[c]	Terrace	%[b]	Detached	Semi-det.[c]	Terrace
Electricity Std & E7	100	2.7	2.1	1.8	100	2.9	2.2	1.9
Coal	100	6.8	5.2	4.4	100	7.0	5.3	4.5
Oil	100	2.7	2.0	1.7	100	2.8	2.2	1.8
LPG	100	1.7	1.3	1.1	100	1.9	1.4	1.2
Gas	100	1.2	0.9	0.8	100	1.4	1.0	0.9

[a] Carbon intensity for each of the displaced fuels are set as given in Table 4-1, with HP systems using the standard electricity. One of the assumptions here is that the dwelling heat demand remains the same regardless of heating system fuel; coal is a further example where this is unlikely to be the case.

[b] This refers to the percentage of ASAPs and GSHPs in cropped sample B2 for which SPFH4 was sufficiently high to result in CO_2 savings.

[c] Based on median gas consumption data from the NEED table creator and 85% efficiency to estimate heat demand for three house types (and allowing for an average of 2% gas used for cooking): Detached 14232 KWh, Semi-detached 10747 kWh, and Terraces 9047 kWh. Sources: https://www.gov.uk/government/statistical-data-sets/need-table-creator

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/386858/Estimates_of_heat_use.pdf

Table 4-3. Median annual fuel cost savings (£ per dwelling) and percentage of HP systems expected to yield energy cost savings relative to other fuels based on cropped Sample B2.

Displaced fuel[a]	%[b]	ASHPs median SPFH4=2.44			GSHP %[b]	GSHPs median SPFH4=2.71		
	ASHP	Detached	Semi-det.[c]	Terrace		Detached	Semi-det.[c]	Terrace
Electricity (Std.)	100	1164	885	749	100	1245	946	801
Electricity (E7)	89	218	166	140	95	298	227	192
Coal	78	126	96	81	90	207	157	133
Oil	2	-202	-153	-130	17	-121	-92	-78
LPG	95	307	233	197	99	387	294	249
Gas	13	-109	-83	-70	35	-28	-21	-18

[a] Fuel costs for each of the displaced fuels are set as given in Table 4-1, with HP systems using the standard electricity tariff. One of the assumptions here is that the dwelling heat demand remains the same regardless of heating system fuel; coal is a further example where this is unlikely to be the case.

[b] This refers to the percentage of ASAPs and GSHPs in cropped sample B2 for which SPFH4 was sufficiently high to result in fuel bill savings.

[c] Based on median gas consumption data from the NEED table creator and 85% efficiency to estimate heat demand for three house types (and allowing for an average of 2% gas used for cooking): Detached 14232 KWh, Semi-detached 10747 kWh, and Terraces 9047 kWh. Sources: https://www.gov.uk/government/statistical-data-sets/need-table-creator

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/386858/Estimates_of_heat_use.pdf

5 Variations in Performance Report

As shown in Figure 3-1 and Table 3-2, sites in the sample show a range of efficiencies. In order to investigate why, the report took two approaches to analysis: statistical and exploratory. First, statistical differences were investigated as a function of heat pump type, heat emitter type and tenure. For this analysis, the largest possible samples were used, consistent with judgements about data quality. A comparison of the performance of two models of air source heat pump was also carried out.

Then an exploratory technical analysis was undertaken with sub-samples of the data used to assess the impact on efficiency of: heat pump cycling, supplementary heating (both domestic hot water immersion and supplementary space heating using the heat pump's internal boost heater, where present), control of domestic hot water immersion and/or boost heating, load factor, and flow temperatures.

The principal results of this analysis are:

a) A wide distribution of seasonal performance factors (SPF) was observed. This appears to be due to both metering errors (of various kinds) and real differences in efficiency, caused by, for example, variations in control and use of resistance heating (immersion heaters etc.).

b) The statistical analysis showed fewer clear results than might be expected. Although GSHPs performed better than ASHPs, and ASHP sites with underfloor heating appeared to perform better than those without, the picture on tenure was more complex. It appears that there are many confounding factors which impact on SPF.

c) Investigation of factors that would be expected to influence performance, such as flow temperature, cycle length and domestic hot water immersion produced the following results:

1. There was no single factor that accounted for good or poor performance.

2. A very large proportion of ASHPs have 10 minute on-to-on cycling patterns. This may be due to the use of boiler thermostats or other ways in which the heat pump controls interact with those in the rest of the heating system. The median on-to-on cycling time of GSHPs was longer, at 18 minutes. Previous lab tests by EA Technology indicated that ASHPs would be expected to show a reduction in efficiency as on-to-off times decreased below 6 minutes. RAPID-HPC's analysis did not show a correlation between median on-to-on cycling period and monthly COP but this lack of relationship may have been influenced by heat metering error.

3. Across the sample, average winter space heating flow temperatures were generally low (<45°C), with only a few sites showing average winter flow temperatures >50°C. Low flow temperatures indicate good design practice and would be expected to result in good efficiencies.

4. During winter, underfloor space heating flow temperatures were lower, on average, than those for systems using radiators. However, there were two underfloor sites with high maximum flow temperatures (>55°C), which would be expected to reduce space heating SPF. These high flow temperatures could be due to a range of contextual technical factors, such as pipes being further apart than recommended, or floor coverings being thick, or underfloor heating being applied to suspended wooden floors, or underfloor heating only occupying a fraction of the floor area of the dwelling, or the presence of mixing circuits between the heat pumps and underfloor heating manifolds[8].

5. Some sites showed excessive use of direct electric immersion heaters for domestic hot water heating, which has a direct adverse effect on SPFH4. On average, where measured, immersion heater electricity consumption was 12% of the total, but more than half of the sites with SPFH4<2 had immersion heating use > 20% of total electricity.

6. There was little evidence of the use of internal boost heating (using internal electric heating cassettes), which would be expected to reduce the heat pump efficiency. This is reassuring; the 2009-2010 EST heat pump field trials showed several examples of excessive use of internal electric cassettes. Note that many designs of heat pump do not contain these cassettes.

7. There were several clear examples of poor control; for example, domestic hot water immersion being used excessively, possibly because of a hot water cylinder thermostat being set to a higher temperature than the heat pump could provide.

---

[8] It should be noted that this is inconsistent with both MCS and industry guidance. But such mixing circuits are standard practice in the context of gas fired central heating, and it is possible that some installers carried the practice across to heat pump installations.

6   Case Studies Report

Heat pump performance is influenced by many factors, including house type, system design and operational patterns. In order to understand the RHPP data better, RAPID-HPC selected 21 properties from the sample for case study analysis.   These properties were located across GB; 14 were owner-occupied properties and 7 were Registered Social Landlord (RSL) properties.

A combined team consisting of an architect and an engineer, with support from a social scientist from RAPID-HPC visited each of these properties to inspect the heat pump installation and monitoring equipment and to interview householders to determine the factors influencing their level of satisfaction.

This report presents the analysis, concentrating on five main areas:

a)   Assessment of quality of the heat pump installation;

b)   Consumer views of heat pumps and examination of the interactions between users' strategies for operation and degree of satisfaction;

c)   Cross-checking of user experience with monitored performance;

d)   Performance of the heat pumps (calculated from measured heat and electricity data);

e)   Exploration of factors influencing performance.

The appendices to the report contain a wealth of data.

6.1   The quality of heat pump installations

For a sub-sample of 10 site visits, the quality of planning and pipe insulation was assessed.   Three heat pump systems were found to be poorly planned/insulated, one was classed as "intermediate" and the remainder were "of good quality". It was not possible to carry out in-depth analysis of radiator sizing and heat demand.

6.2   Consumer views of heat pump systems

The team interviewed householders to establish their strategies for operating their heat pump, their estimated bills, the degree of comfort, how the heat pump compared with their previous fuel and whether they had experienced any problems with their heat pump.

In eighteen out of twenty-one cases, occupants were satisfied or very satisfied with their heat pumps and preferred them to their previous heating system. But the case studies revealed the rich complexity of the

notion of satisfaction, which included the level of thermal comfort felt, running costs, ease of use, environmental impact, technical integrity, noise levels and controllability of the system.
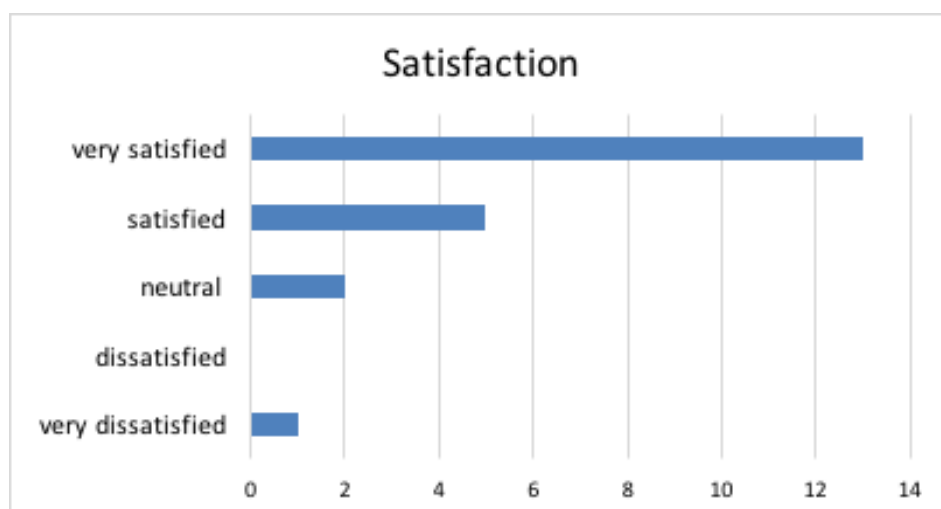

Figure 6-1. Levels of satisfaction reported across the 21 case studies.

Reasons stated for satisfaction varied from household to household and any given household's overall satisfaction level was generally a synthesis of several different factors (for example, cost and constant heat (CS20) or maintenance and environmental benefits (CS14)). One social housing case stood out; the occupants were dissatisfied for a variety of reasons that did not seem to correlate with the apparently good performance of the heat pump during the final year of monitoring.

Satisfaction with the training material provided was a little lower, with 17 of the 21 households stating that they were satisfied or very satisfied with the training material provided.

6.3   Consumers' strategies for control

Householders were asked about their strategies for controlling their heat pump, including what their thermostat settings were; whether they had zonal control or radiator TRV's; whether they operated the system 24/7 or for shorter periods; whether they used night setback temperatures; how much domestic hot water they used and whether they used secondary heating (for example log fires or electric heaters).

The results indicated a widespread of patterns of use. For example, thermostat set points generally ranged from 18 to 23°C (with one exception for a householder with health problems). It appeared that the social housing tenants had been asked to control their heat pump by the thermostat only. Several private householders deliberately experimented with different operational strategies. One of these stated that he eventually settled on the simplest strategy, of continuous heating.

6.4   Faults

Despite the high levels of householder satisfaction, at least 10 out of the 21 cases had experienced some significant problem since installation. Issues as described by occupants included faulty HPs or faulty sub-system, installation and antifreeze problems, condensation dripping from external ASHP units, blockages,

a "faulty motherboard", and unintentional use of resistance heating resulting in excessively high electricity bills.

Despite significant disruptions (HPs in three out of the seven RSL cases suffered a major breakdown, with heating systems out of action for periods of up to two months), six out of seven RSL households were satisfied with their new heating systems. Responses from RSL occupants show the need for RSLs to have access to competent personnel to deal with troubleshooting.

6.5   Cross-checking of site data with monitored data

The site visits identified a number of errors in the metadata, including inaccurate entries in the MCS certificates, inaccuracies in information about emitters and inaccuracies in EPCs. Unfortunately problems with metadata cannot be detected solely from statistical analysis of the larger dataset. This finding emphasises the importance of ensuring that resources are available for site visits in field trials.

Interview data were cross-checked against the monitored heat and electricity data. In one case, a householder stated that the heat pump settings had reset after a power cut and that subsequently bills had increased. Examination of the data indicated that the domestic hot water immersion had been used excessively during this period.

In another case, the householder was adamant that the performance was good, despite the apparent poor performance of the heat pump. Detailed examination of the heat and electricity data indicated that there had been a problem with heat metering during the period in question, and that selection of a different period yielded a significantly higher estimate of SPF.

As a result of detailed examination of the heat and electricity data and cross checking with interview data, three of the twenty-one estimates of seasonal performance factor (SPFH4) were revised upwards.

These examples demonstrate the value of combining social and technical information.

6.6   Calculated performance of heat pumps in the case study sample

The distribution of SPFH4 in the 21 case study sites is shown in Figure 6-2 and shows a large range of performance.
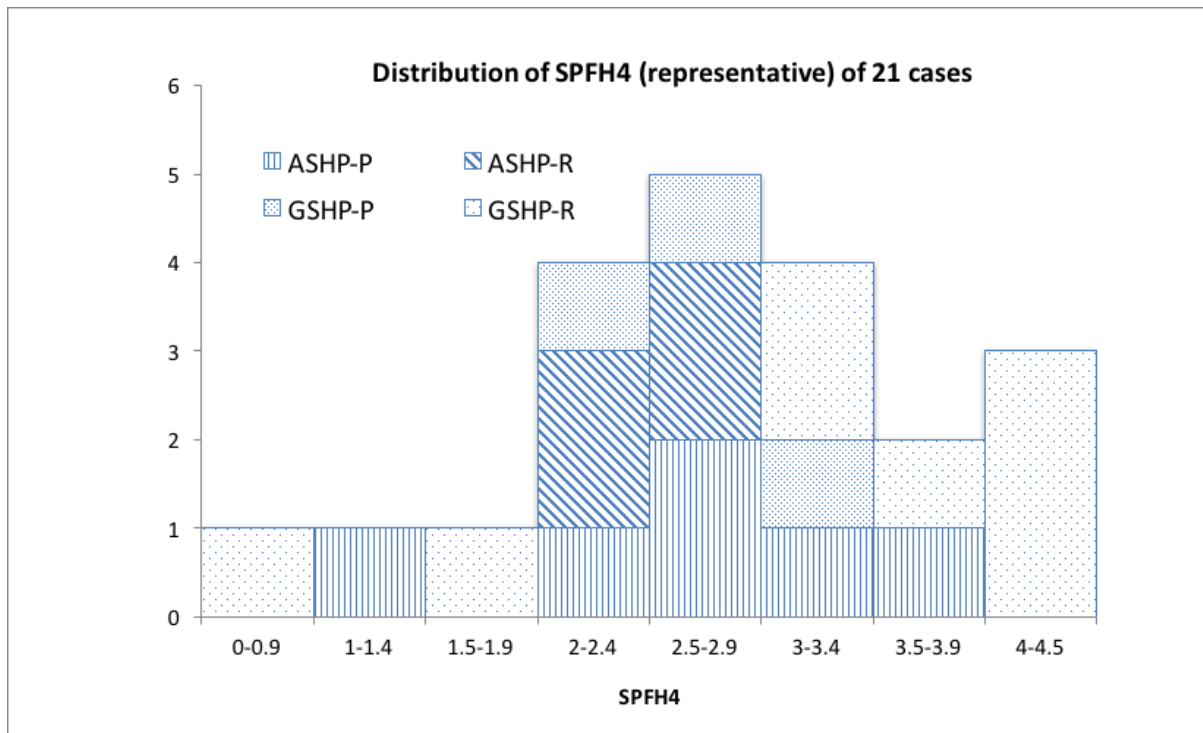
Figure 6-2 Distribution of representative SPFH4 across the 21 case studies by heat pump type and tenure (P="Private", R="Rented from RSL") .

6.7 Factors influencing performance

The study indicated that a range of social and technical factors with the potential to influence performance:

- Comparison of performance of three cases with similar GSHPs in similar sized dwellings, strongly suggested that SPF was sensitive to load factor for these systems. Low load factors can arise in several ways including mis-sizing of the heat pump, and excessive use of supplementary sources of heat.

- Excessive use of electric resistance heating (particularly domestic hot water or internal boost being switched on unnecessarily).

- In one case, location of the heat pump far from the house to minimise noise issues, would have resulted in loss of heat in the pipes between the heat pump and the house.

These issues, though at first sight physical in nature, are the result of both physical factors (e.g. system design, installation, the way controls are set up) and social factors (e.g. how designers and installers interpreted and responded to requirements of householders, lifestyles and how occupants chose to control their heat pumps).

Experience on a number of sites suggests that follow up visits by technically competent personnel might help to ensure that initial teething issues were resolved and that performance was maintained over the long term.

7   MCS Report

The Microgeneration Certification Scheme Installation Standard (MCS MIS) 3005 provides the 'requirements for contractors undertaking the supply, design, installation, set to work commissioning and handover' of microgeneration heat pump systems for compliance with the certification scheme.

The aim of this report was to assess how well RHPP Trial installations reflect the design requirements of MCS MIS 3005.

The MCS MIS 3005 standard was changed several times during the period over which RHPP heat pumps were designed and installed. Some of these changes were significant. In principle, systems should have been designed to whichever version of the standard was mandatory at the time of *quotation*, not installation. The metadata supplied with the RHPP data does not include the version number of MIS 3005 used for the design, and quotation dates are only provided for privately owned properties. It is therefore not possible to determine with certainty which version of the standards was applied by the designer in each case.

For these reasons, the assessment in this report cannot provide precise estimates of how many systems complied with the standards. However, eight elements of the MIS 3005 standard were examined in detail, namely:

- Calculation of heat loss
- Heat pump sizing
- Radiator sizing
- Calculation of measured annual energy use and comparison with the installers' estimates and EPC calculations
- Sterilisation of domestic hot water
- Specification of flow temperature at design conditions.
- Weather compensation
- Actual measured SPF_{H2} for space heating compared to the SPF_{H2} predicted from the MCS Heat Emitter Guide (MCS 021, MCS, 2015[9]).

---

[9] The "Likely SPF_{H2}" values are not provided in version 2.1 of the MCS "Heat Emitter Guide" but can be found in the archived version 2.0:
http://www.microgenerationcertification.org/images/MIS_3005_Supplementary_Information_2_-_Heat_Emitter_Guide_v2.0.pdf

Data taken between 1/11/2011 to 31/10/2015 was used in the analysis. A subset of the data for just one year (the concurrent dataset) between 1/11/2013 and 31/10/2014 was also used when appropriate[10].

7.1   Building heat loss calculations

Heat loss calculations are necessary for the selection of appropriately sized heat pumps (or any other heating appliance), whether in new build or retrofit. Heat loss calculations may currently be carried out by hand, spreadsheet or using proprietary software. In order to support installers, MCS introduced an approved heat loss calculator on 3rd November 2015.

The installer's estimate of heat loss for the building affects both the estimate of annual electricity use and the sizing of the heat pump. Undersizing potentially increases the likelihood of electric resistance heating, while, as discussed above, oversizing may result in low load factors. Both have the capacity to reduce SPF.

It was not possible to verify the installers' estimates of heat loss for the properties in the RHPP sample, since design criteria such as areas and volumes, thermal transmittance values (U-values), allowances for ventilation, etc., were not provided in the metadata.

However, the report gives examples of heat loss calculations for a single property, and demonstrates that estimated heat loss can be strongly influenced by subjective assessment of U-values and ventilation rates, in particular for retrofit situations. For example, calculations for an end-terraced house showed a 24% increase in estimated heat loss when ventilation assumptions were changed from ventilation class C to ventilation class A, as defined in BS EN 12831. It should be noted that this issue is not specific to heat pumps, but heat pumps may be more sensitive to errors and uncertainties around sizing because:

- HP performance appears to be more strongly affected by part loading than modern gas boilers[11];

- the marginal cost of heating capacity (£ per marginal kW of heat output) is significantly higher for heat pumps than for gas boilers.

7.2   Heat pump sizing

Heat pump sizing for design conditions was found to be either poorly understood or poorly expressed. At the time of the trials, installers predominantly assessed 'net capacity' as manufacturers' nominal capacity and not at the site specific design conditions. RAPID-HPC used the measured heat output from the RHPP sample to extrapolate to a daily external temperature of 0°C and compared this estimated power with the nominal capacity as quoted. Note that the choice of a daily average temperature of 0°C

---

[10] Further information on the process of selecting data for analysis is given in the _RHPP Performance Variations Report_.
[11] It should be noted that the RHPP programme was not established to support replacement of gas boilers with heat pumps. But these comparisons are relevant in a discussion of the standards for calculating heat loads for houses for the purpose of sizing new heating equipment.

was a simplification[12]. Manufacturers' data was used to provide estimates of the difference between nominal capacity and capacity at design conditions for a range of heat pump types. Comparing two extrapolated figures is necessarily subject to error. However, a majority of heat pumps appear to be adequately sized when compared to peak measured load.

As mentioned, heat pump sizing is affected by the calculated heat loss, which is sensitive to assumptions in ventilation and U-values. The same issues apply for boiler sizing, but the cost and efficiency implications of over-sizing boilers are lower.

7.3 Comparison of installers' estimate of annual heat demand with measured values

The installers' estimates of annual energy use are influenced primarily by five factors: weather, estimated heat loss and assumed SPF of the heat pump at the chosen flow temperature, proportion of space heating relative to water heating and the assumed SPF for space and water heating.

A comparison of installers' estimates of annual heat demand with measured values indicates a relatively poor correlation. Calculating annual energy use by integrating the technical complexities of heat pump whole-system performance and occupant operational preference is a complex socio-technical challenge.

The comparison of kWh/year from Energy Performance Certificates (EPC) with trial data also showed a discrepancy, but the most significant source of disagreement may be prevailing, wide variations in commercial EPC ratings of dwellings (DECC, 2014a).

The poor agreement between measured and estimated energy use may be due to mild winters during the trial, or may suggest that calculation procedures were too complex.

7.4 Radiator Sizing

Radiator sizing analysis indicates that 'star rating oversize factors' as described in the Heat Emitter Guide (HEG) may be inadequately understood or ignored due to practical and aesthetic considerations of size and location.

7.5 Sterilisation of domestic hot water

There is no clear understanding of the number of systems installed after compliance with MIS 3005 v3.1a was made compulsory from 1st March 2012. Metadata quotation dates are given only for private housing and installation/commissioning dates from MCS certificates are only broadly indicative of the period of

---

[12] Ideally, the extrapolation would have been made to the design temperature for each heat pump. A daily average of 0°C was judged to be sufficiently cold to comprise hours for which the temperature was equal to, or lower than, the design temperature of the heat pump. The y-axis of the charts referred to the 99th percentile of the flow temperatures; it is reasonable to assume that these would have occurred during the coldest hours of the day.

design and installation since, for a domestic heating installation, the installer's quotation may precede *or* be followed by a full technical specification compliant with MIS 3005.

Compliance with protection from legionella exemplifies this uncertainty. Those systems quoted for after version 3.1a became mandatory should have included appropriate measures to ensure protection against legionella. For those RHPP heat pumps capable of producing water at an appropriate temperature to achieve cylinder temperatures of 60°C or above, sterilisation is unnecessary although immersion heaters may be still be present for back up purposes.

Where normal cylinder storage temperatures are below 60°C, the installation should have incorporated regular sterilisation of the domestic hot water. Whilst cylinder hot water temperatures were not monitored, examination of the data from 220 metered immersion heaters in the sample indicates that between one quarter and one third of these exhibit immersion operation consistent with regular sterilisation; predominantly either weekly or daily, although other patterns also emerged.

7.6   Design Flow Temperatures

The analysis of maximum flow temperatures at minimum outdoor design temperature indicates a mean of between 40 and 45°C for both radiator and underfloor heating corresponding to 4 star operation from the Heat Emitter Guide (HEG) (MCS, 2015). This would be expected to result in good performance.

However, a wide range of temperatures is observed; 17% of systems examined had design flow temperatures of 50-60ºC, indicating 2 or 1 star operation in the HEG, and 34% had design flow temperatures of <40ºC, indicating 5 or 6 star operation. As noted earlier, extrapolation using regression calculations down to 0°C was necessary due to the low number of days in the RHPP dataset with daily mean external temperatures of 0°C.

7.7   Weather Compensation

The same analysis indicates weather compensation was used for 64% of installations. Weather compensation is recommended by MCS MIS 3005 version 4.0; however, under some circumstances, for example, intermittent heating, weather compensation may not be the most effective strategy.

7.8   Comparison between measured space heating SPF and Heat Emitter Guide "likely space heating SPF"

Using the design flow temperatures calculated for each site, the Heat Emitter Guide (HEG) 'likely space heating SPF' has been compared to the actual, measured space heating SPF. Correlations are poor, with the observed SPF's being significantly lower than the HEG values. This is more pronounced for GSHP than ASHP.

In conclusion, MCS heat pump installation standards were updated significantly and on several occasions during the RHPP period. Any changes inevitably take time to embed. The analysis presented here refers to the monitored RHPP sample; it should not be assumed to apply to heat pumps installed after the RHPP.

8   Overall conclusions

The analysis presented here demonstrates that median SPFH2 values for ASHP and GSHP in the cropped B2 sample are around 2.65 and 2.81 respectively. Approximately two thirds of ASHP and four fifths of GSHP in the cropped B2 sample met the EU criterion for being considered "renewable". Filtering the data in alternative but equally plausible ways results in similar conclusions.

A detailed assessment of monitoring errors has been undertaken. Some of these errors increase SPF, others decrease it. Most only affect a few sites and only two have a significant effect on the medians across the sample:

- Missing heat data, when electricity data are present – applies to some heat pumps and could cause the median SPFH2 to be *under-estimated* by <4%, although individual sites could be more affected.

- Heat meters not calibrated for glycol – is likely to apply to most heat pumps and causes the median SPFH2 to be *over-estimated* by 4-7%.

It is estimated that 89% and 95% of the ASHPs and 95% and 99% of GSHPs in the cropped B2 sample would have saved on bills relative to electric storage heating and LPG respectively. Estimates of bills savings are sensitive to fuel prices and tariffs. Based on the tariffs assumed here, estimates of bill savings relative to oil and gas are negative. Prices of oil and gas have fallen in recent years; if prices from 2011-2014 are used, then the calculation shows a bill saving for heat pumps against oil.

Heat pump performance is sensitive to context of dwelling, including the heating system, controls, commissioning, and operational and lifestyle decisions of occupants. The case studies revealed the complexity of the notion of satisfaction, which included the level of thermal comfort felt, running costs, ease of use, environmental impact, technical integrity, noise levels and controllability of the system. 18 out of the 21 householders described themselves as "satisfied" or "very satisfied" with their heat pump.

Controls should be optimised to ensure that resistance heating (especially for domestic hot water) is not used excessively, whilst still ensuring regular sterilisation.

The RHPP study represents an important contribution to the understanding of domestic heat pumps in the UK. The study has demonstrated that combining technical monitoring with site visits to assess installations and with householder interviews is a very valuable way to improve understanding of heat pumps. Further recommendations for future field trials are given in section 9.

9   Guidance on future field trials

9.1   Study design principles and rationale

Apart from cost, the major consideration in the design of a field trial is to identify the key objectives of the project since this will frame the design specifications, for instance in terms of the size of the study and the types of data that are collected. For future heat pump field trials, these objectives could be as follows (progressing from the more general outcomes to more specific or detailed findings):

1. Provide an overall assessment of performance (e.g. SPFs) of the existing installation base of heat pumps - or at least of a cohort of heat pumps installed between specific dates related to the study and therefore reflecting a generation of technology.

2. Quantify the energy saved (and reduced carbon emissions) through the use of HP systems, allowing for changes in occupant behaviour including the possibility of higher indoor temperatures consequent upon a shift from intermittent to continuous heating.

3. Detect progress or change over time, for example with respect to heat pump performance that corresponds with external changes such as improvements in heat pump technology and publication of improved installation guidelines.

4. Identify/detect performance/operational issues that may occur, for instance with respect to a particular type of HP system or with installation on a specific type of site.

5. Advance understanding of some of the underlying/deeper factors at work that may impact adoption rates or heat pump usage and heating patterns.

Each of these objectives carries implications for the study design. Objective 1 requires a large sample size selected with sufficient heterogeneity to cover key characteristics of both heat pumps and sites (dwelling type and occupant characteristics). Note that heterogeneity of the sample characteristics is more important than representativeness. It is inefficient to try and have every group in the same proportion as they appear in the population or stock of buildings. A better strategy is to oversample to ensure sufficient sample size within the smaller subgroups of interest, and then weight the overall sample according to their prevalence in the population when generating estimates of stock summary statistics. This is the strategy adopted by the Survey of English Housing and other major UK social surveys. So if a group of sites of particular interest constitute only 5% of the stock or population, sampling them so that they are 10% of the sample (e.g. so that there are at least 30 in the subsample) means the overall sample size can be reduced.

It should also be considered how precise such sub-group statistics need to be in terms of the implications for policy. A much larger sample and more rigorous selection process (to reduce sample bias) would be needed if tight quantification of performance is needed, whereas it may be possible to relax these criteria somewhat if only a broad categorisation of heat pump performance were required, for example into categories such as *poor*, *moderate*, and *good* performance. Broad groupings might still serve to trigger a policy response if too high a proportion of sites were in the poor and moderate groups, without needing to know performance figures with precision. Similarly a policy success may be identified by showing that vast majority of sites are in the *good* performance category. To reiterate, these types of qualitative findings may be more significant in terms of policy development than having a statistically narrow estimate of average performance of HP systems across the sample, while reducing the size and cost of the field trial. Statistical power calculation provide a means to estimate the sample size needed to detect a difference in average performance between two groups. This needs to be considered in light of likely inherent variability of estimates of heat pump performance over a year, given the range of likely factors affecting *in-situ* performance over a year, even with perfect metering.

Both Objectives 2 and 3, imply measurement before and after the installation or other change of interest occurred. An alternative, referred to as a case-control study, would involve having a control group, for instance who did not receive a heat pump, typically three or more sites matched to each participating site so that some comparison can be made. However it is likely to be more practical to instead rely on having sufficient data from before and after the intervention (installation of a heat pump). This potentially impacts the duration of study, unless research teams have reliable access to energy bill data for a reasonable period preceding the installation of new heating technology. If SPF, which is defined as an annual measure, is the primary means of benchmarking energy performance of the heat pump, then at least one year of measurements would be required prior to the installation and after the event of interest. In practice, this leads to conclusion that sites which already have smart meters should be recruited (which in turn implies gas or electric sites, though it might be possible to estimate oil and solid fuel consumption from deliveries) for at least one year, and seeking permission of the household to obtain metering/billing data. It also raises the question that SPF may not be the most appropriate metric to use, unless performance measurement over a shorter period can be shown to be robustly converted to annual assessment (including accounting for variations in performance in relation to external temperature).

Both Objectives 4 and 5, have a range of implications for the set of measurements, and the degree of data collection around the site and from the occupant. There may be important information from this work that does not require waiting for completion of a full year of measurements for it to have wider relevance, including on the survey design; it is highly unlikely to be practical to undertake this level of in-depth monitoring and survey work needed in these sites for the entire sample. For this reason case studies can form a key part of the strategy to develop and inform the broader field trial programme.

The value of case studies lies in the close examination and detailed understanding of individual cases – a case study is not a small field trial. The flexible nature of case study methodology lays the research team

open to the unexpected to a greater extent than other approaches. In the present study it enabled issues – such as the effect of load factors on SPF, the complex nature of occupant satisfaction and the presence of a gas supply in four of the case study dwellings - to emerge in ways that would have been unlikely or impossible through statistical analysis of data collected by remote physical monitoring alone.

A critical decision in the design of field trials is the number of cases that are to be included. This must be considered together with the issues that are to be explored and the resource that is available. One of the mistakes often made by researchers new to case studies, is to take on too many cases, sometimes with the aim of undertaking statistical analysis within the sample of cases in pursuit of a narrow interpretation of generalisability.

Much of the work on the case study conducted with the RHPP sample was undertaken by an architect and a heating system engineer, with support of an experienced social scientist. But it also required other members of the consortium, whose main focus was analysis of monitoring data, to provide supporting data and analysis. The resource implications of such support were underestimated in this study. In this example the research team set out to study twenty cases, with one additional case undertaken as a pilot. But resource limitations meant that detailed analysis was initially undertaken on a sub-set of just ten cases.

In the context of case studies on the deployment of energy systems, greatest benefit is likely to be gained from a socio-technical approach (Chiu et al., 2014) using mixed methods (Gray, 2004). It is critical that case studies in this field are not conflated with the purely social and purely qualitative.

Case studies practised in this way can support large scale, predominantly quantitative, statistically oriented field trials in at least two ways. If undertaken during the installation of monitoring equipment, they can help to improve the quality of both metadata and data collected by remote monitoring systems. If, as in the study reported here, they are undertaken after the onset of physical monitoring, they can make use of insights from monitoring data to suggest lines of enquiry for site visits and subsequent analysis. In this study, the value of such insights was limited by the quality of monitoring data available. High quality monitoring data is as important to mixed methods case studies as it is for the conduct of large scale, predominantly quantitative, statistically oriented field trials of new energy technologies.

As an example of the kind of issues that one of the lines of inquiry of this case study was to explore the relationships between dwelling heat load, heat transfer capacity of the heat distribution system, and heat pump capacity. It transpired that to fully achieve this would have required significantly more resource and time on site than was available. There is a possibility that longer site visits would affect the number of households that would be willing to engage with a future research team. It is possible that this could be mitigated by requiring recipients of subsidies for new technologies to indicate their willingness to engage with a future research project, should one be commissioned. This should include the option of additional contact with dwelling occupants to resolve questions that emerge after an initial site visit.

The above lessons for case studies can be summarised as follows:

- socio-technical, mixed methods case studies are likely to add significant value to any field trial of new energy technologies;

- depending on the design of the field trial, and specific objectives case studies can be undertaken either during the installation of monitoring equipment, or after the onset of physical monitoring;

- a third, more ambitious option would be to undertake longitudinal case studies, which followed individual cases from recruitment into a programme, through design, installation and occupation – if adequately resourced such an approach has the potential to provide very rich insight into the multiple factors at work in the performance of energy technologies;

- case studies should examine numbers of cases appropriate to research objectives and available resources – except in very unusual circumstances, no more than twenty;

- the value of case studies is likely to be significantly enhanced by the availability of high quality physical monitoring and energy bill data.

9.2　Study design description

Based on the above considerations, it is suggested that a stratified approach should be taken:

Level 1 Study: the baseline data collection: A large sample of sites (>1000) should be recruited that intend to install or have recently installed heat pump systems, with the addition criteria that they have given permission for the researchers to obtain data from their smart meters (or fuel supply bills if oil or other liquid/solid fuels) regarding their total energy use for at least the previous year. A high proportion of sites will only need to provide this information and have a reduced SAP survey and basic occupant survey done (e.g. at the time of recruitment). They form the frame for the other studies in the field trial and described below. As noted above this sample should be sufficiently heterogeneous to cover the groups of interest, such as by size and heat pump and dwelling and occupant characteristics.

Level 3 Study: detailed case study: based on the type of heat pump system, site characteristics, and household, a detailed survey and monitoring programme (informed by the previous work) should be undertaken, potentially in stages as resources permit and in order to expand the range of sites progressively as the methodology is resolved. Specifically, this sociotechnical investigation study should trial a series of revisions and refinements to the data collection methodology, especially in regard to the metering equipment and the occupant and dwelling surveys. Based on short term assessment of heat pump performance, this should continue until such time that metering issues have been resolved to all reasonable satisfaction, and a reliable protocol can be developed for a range of installation. The sample size should progressively expand to around 50 sites, or until the methodology is deemed sufficiently stable and reliable for wider application. The methodology should include development of a series of checks or tests (e.g. on a monthly basis) to ensure the metering system is working satisfactorily and that anomalies reflect operational issues. It should be expected that sites may need repeated visits to identify and resolve

issues. This study is in effect an intervention study, since issues with the heat pump itself will likely be identified and fixed as part of resolving issues with metering. Thus the performance results, which as a consequence one would expect to be better than usual, should not be aggregated with those from other the other studies. Note that by drawing on the insights gained from the existing RHPP study, including the case study, a smaller sample size may be sufficient to develop the methodology sufficiently to justify proceeding with the next stage of the research programme.

Level 2: the main study: select the study sample from level 1 sites, such that across each subgroup of interest, for instance GSHPs with underfloor heating in private homes with more than three bedrooms, there should be a minimum of n=12 sites. Any less than this and the subgroup will not have size to enable a meaningful comparison with other groups. Important subgroups that represent major categories of the population of sites will need to be considerably larger, and are typically size on their prevalence. This sample of sites have the full monitoring set-up installed, as well as additional site and occupant surveys. The insights and strength of the methodology developed from the Level 3 study, should provide some insurance that the retention rate remains high in this sample. The monitoring data should be updated continually through the course of the study. Using the regular tests developed previously, if evidence of an unexpected pattern in the data emerges across sites, then a few of them can be approached for more detail investigation (and in effect they will join the Level 3 case study sample).

One of the key advantages of the level one study, which has an independent measurement of delivered energy (from the smart meter or billing data) is that it provides an independent check on changes in energy demand before and after installation. This has two main implications. It provides a 'sense-check' against the monitoring performance of the heat pump. Second, these measures provide an overall estimate for the change in delivered energy due to the installation of heat pump systems, including any related changes in heating patterns on indoor temperatures in response to the presence of the heat pumps or subsequent lower energy bills. In other words this would provide empirical evidence of the impact of heat pumps, including possible increases in annual demand for heat, which to date have only been approximately estimated.

References

BS EN 12831:2003, *Heating systems in buildings. Method for calculation of the design heat load*, London: BSI.

Chiu, L.F., Lowe, R., Raslan, R., Altamirano-Medina, H. & Wingfield, J., 2014. A socio-technical approach to post occupancy evaluation: interactive adaptability in domestic retrofit, *Building Research & Information*, 42 (5) 574-590. http://dx.doi.org/10.1080/09613218.2014.912539 .

DECC, 2014a. *Green Deal Assessment Mystery Shopping Research: Mystery shopping of customer experiences of Green Deal assessments with 48 households and analysis of variability across Green Deal Assessment Reports of 29 properties*, London: DECC.
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/388197/Green_Deal_Assessment_Mystery_Shopping_FINAL_PUBLISHED.pdf.

Dunbabin, P., & Wickins, C., 2012. Detailed analysis from the first phase of the Energy Saving Trust's heat pump field trial: Evidence to support the revision of the MCS Installer Standard MIS 3005 Issue 3.1. London: DECC.
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/48327/5045-heat-pump-field-trials.pdf

Dunbabin, P., Charlick, H., & Green, R., 2013. Detailed analysis from the second phase of the Energy Saving Trust's heat pump field trial. London: DECC.
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/225825/analysis_data_second_phase_est_heat_pump_field_trials.pdf

European Commission, 2013. *Establishing the guidelines for Member States on calculating renewable energy from heat pumps from different heat pump technologies pursuant to Article 5 of Directive 2009/28/EC of the European Parliament and of the Council.*

Gleeson and Lowe, 2013. Meta-analysis of European heat pump field trial efficiencies. *Energy and Buildings*, 66: 637-647.

Gray, D., 2004. *Doing Research in the Real World*, Thousand Oaks: Sage.

Green, R., 2012. The Effect of Cycling on Heat Pump Performance. London: DECC.

H.M. Treasury, 2015. *Green Book supplementary guidance: valuation of energy use and greenhouse gas emissions for appraisal*, London: Treasury, https://www.gov.uk/government/publications/valuation-of-energy-use-and-greenhouse-gas-emissions-for-appraisal .

Hughes, D., 2016. Monitoring of Non-Domestic Renewable Heat Incentive Ground-Source and Water-Source Heat Pumps. London: DECC.

Miara, M., Günther, D., Kramer, T., Oltersdorf, T., Wapler, J., 2011. *Wärmepumpen Effizienz - Messtechnische Untersuchung von Wärmepumpenanlagen zur Analyse und Bewertung der Effizienz im realen Betrieb*, Fraunhofer-Institut für Solare Energiesysteme ISE, Freiburg. (Heat pump efficiency - measurement study of heat pump systems for the analysis and assessment of the effectiveness in the real operation). Fraunhofer-Institute for Solar Energy Systems. http://wp-effizienz.ise.fraunhofer.de/download/wp_effizienz_endbericht_kurzfassung.pdf

MCS 021 2015 *"Heat Emitter Guide For Domestic Heat Pumps"*, *Issue 2.1*, London, Microgeneration Certification Scheme.

Pedersen S.V. & Jacobsen E., 2011. *Approval Of Systems Entitled To Subsidies. Measurements, Data Collection And Dissemination*, Taastrup (Denmark): Danish Technological Institute.

RAPID-HPC, 2016. *Detailed analysis of data from heat pumps installed via the Renewable Heat Premium Payment Scheme (Superseded)*, London: DECC. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/499194/DECC_RHPP_160112_Detailed_analysis_report.pdf

RAPID-HPC, 2017a. *Investigating variations in performance of heat pumps installed via the Renewable Heat Premium Payment Scheme*. London: BEIS.

RAPID-HPC, 2017b. *RHPP report on compliance with MCS installation standards*, London: BEIS.

RAPID-HPC, 2017c. *Case Studies Report from the RHPP Heat Pump Monitoring Campaign*, London: BEIS.

RAPID-HPC, 2017e. *RHPP - Note on Systematic Errors in Physical Monitoring Data* (reissued with minor corrections, originally issued 2016), London: BEIS.

Riviere, A., Coevoet, M., Tran, C., Zottl, A., Nordman, R., 2011. D4.2. / D 3. 4 . *Concept for evaluation of SPF - Version 2.0 - A defined methodology for calculation of the seasonal performance factor and a definition which devices of the system have to be included in this calculation: Air to air heat pumps*. https://ec.europa.eu/energy/intelligent/projects/sites/iee-projects/files/projects/documents/sepemo-build_concept_for_evaluation_of_spf_for_hydronic_hps_en.pdf

Tukey, J., 1977. *Exploratory Data Analysis*, Addison-Wesley.

Wickins, C., 2014. Preliminary report on the Renewable Heat Premium Payment metering programme. London: DECC.
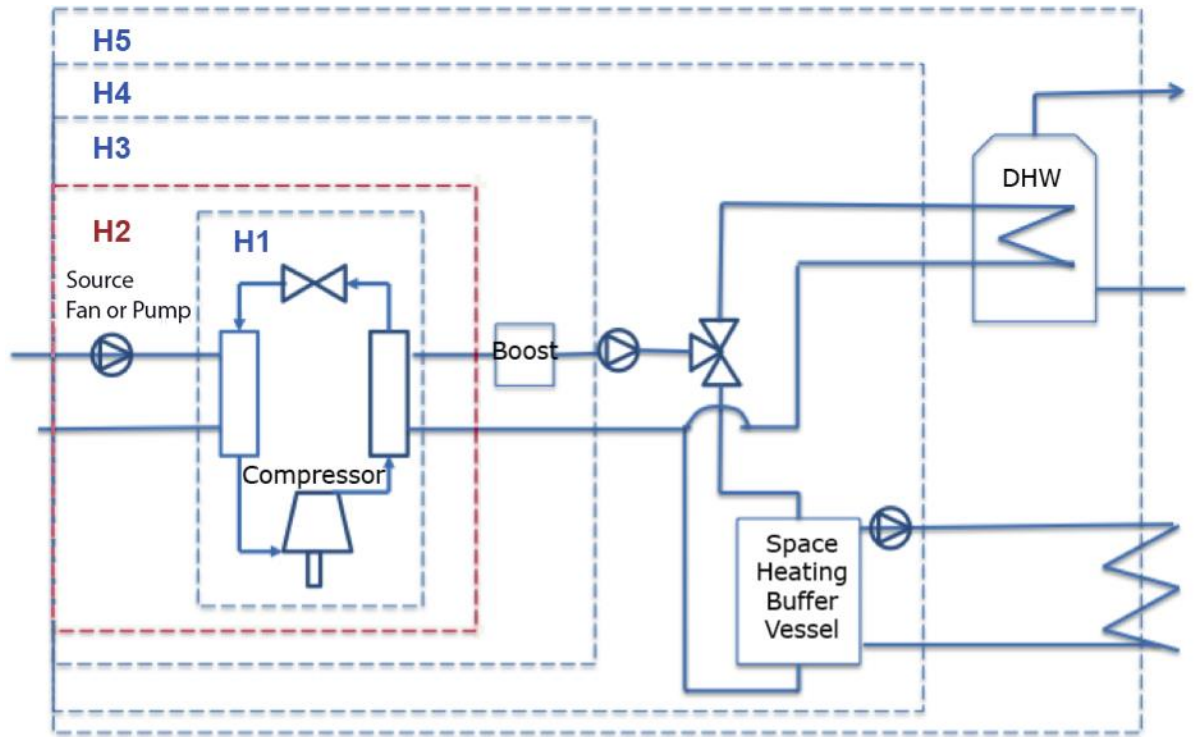
Appendix A1

Figure A 1 SEPEMO system boundaries (derived from Riviere et al., 2011) with the addition of H5 boundary that accounts for heat losses from the hot water cylinder.

Appendix A2

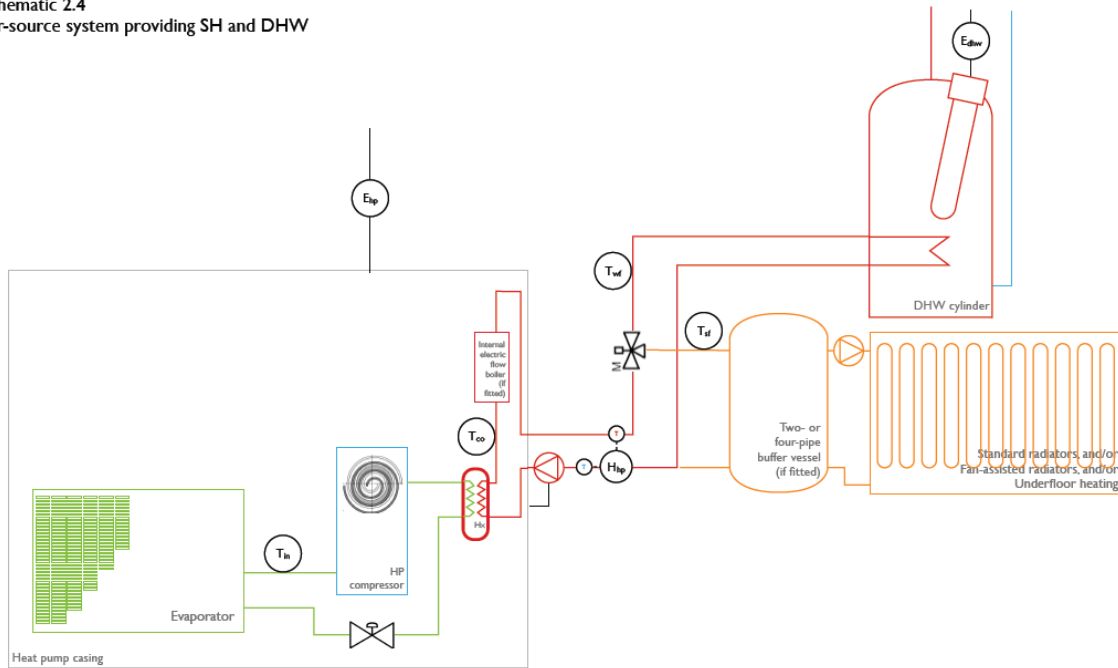Schematic 2.4
Air-source system providing SH and DHW

Figure A 2. An example simplified schematic of the metering arrangement for a monobloc ASHP that provides heat to space heating and a domestic hot water cylinder with an immersion element.
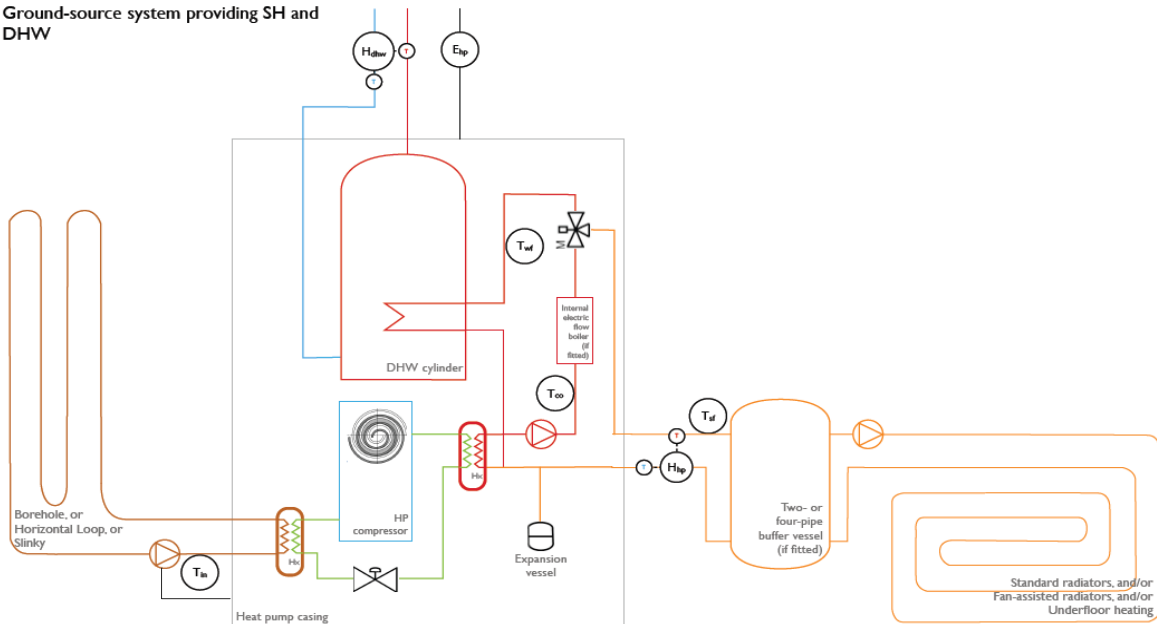
Schematic 16.12
Ground-source system providing SH and DHW

Figure A 3. An example of a GSHP with an integrated domestic hot water cylinder

Appendix A3

Meter type	Potential Fault type	Description	How do we know that these faults exist?	Systematic error or an error that affects individual sites?	Effect on SPF or monthly COPs
Heat meters	Missing heat meter data, when electricity data is present.	Periods with zero or unusually low heat data were not filtered out in the data cleaning process.	Observed in data	Individual sites	Will have the effect of under-estimating SPF, by an estimated ~4% across the Sample B2 (cropped) but much higher for a few sites. Apparent slight effect on distribution of SPFs (statistical tests not carried out to confirm this).
Heat meters	Systematic under-reading due to meter installation.	Poor installation of strap-on sensors or pocket sensors RAPID-HPC removed 99 sites with known strap-on sensors at the start of the project, but suspect that others may exist.	Some suspiciously low COP readings observed in data (e.g. < 1)	Individual sites	Would reduce SPF and monthly COP but sites for which spfh4<1.5 have been filtered out of Sample B2 (cropped).

Heat meters	Systematic over-reading due to glycol correction not being applied.	Heat meters calibrated for water with no antifreeze.	Wickins (2014)	Likely to occur in many of the sites, in both ASHPs and GSHPs.	Likely to result in over-estimation of SPF by 4-7% - see separate report on systematic errors.
Heat meters	Limited to 18 kW	Up to 16 sites in Sample B2 (cropped) affected.	Observed in data	Individual sites	Expected to affect the SPFs of these sites slightly in cold weather.
Heat meters	Systematic over-reporting of heat output.	Probably due to heat meter temperature sensor offsets, exacerbated by circulation pump over-run.	Observed in data (heat output when no electricity input)	Individual sites	Will over-report SPF and COP.
Heat meters	Spikes in heat output when changing mode.	Not known whether this is a metering problem or a real dynamic effect with no impact on estimates of heat.	Observed in data.	Individual sites. Not present in all sites, but for those in which this effect is present, it occurs every time there is a mode change.	Unknown. If real heat, no effect, if metering error, over-reports SPF and COP.
Heat meters	Transposition of Hhp and Hhw sensors.		Observed in data.	Individual sites	No effect on overall SPF or COP, but will affect space heating and DHW SPFs and COPs.
Heat meters	Flow decay over the dataset time period. Median 1.5% decay over year for Sample	Cause unknown.	Observed in data	Individual sites	Under-report SPF, and COP for later months.

	B2 (cropped).				
Heat and electricity meters	Heat and electricity data missing at the same time.	Cause unknown.	Observed in data; of 34 sites investigated in detail, 16 had > 7 days of this.	Individual sites	Effect depends on the time of year at which the problem occurs.
Electricity meters	Suspected unmetered electricity – missing Eboost or Edhw	Temperature data shows unusual patterns which can't be explained by the existing heat and electricity data.	Observed in data	Individual sites	Over-report SPF.
Electricity meters	Transposition of electricity meters.	Transposition of Ehp and Edhw or Eboost, or Edhw and Eboost.	Observed in data. Automatic correction applied in code for cases where easily detectable but not all cases.	Individual sites	Effect depends on which sensors were involved. Overall SPFH4 unchanged but other boundaries affected. Space heating and DHW SPFs could also be affected.
Temperature	Sensors too close to other pipes.	This causes e.g. Tsf to be influenced by Twf and vice versa. This in turn affects which mode (space heating, DHW) gets attributed to each 2 minutes of data.	Observed in data and photos	Individual sites	Overall SPF and COPs are unaffected but space heating and DHW SPFs and COPs are affected.