



Animal &
Plant Health
Agency

Livestock
Demographic
Data Group:
Pig
population
report

Livestock
population
density maps
for GB

November 2017





© Crown copyright 2017

You may re-use this information (excluding logos) free of charge in any format or medium, under the terms of the Open Government Licence v.3. To view this licence visit www.nationalarchives.gov.uk/doc/open-government-licence/version/3/ or email PSI@nationalarchives.gsi.gov.uk

Document information	
LDDG	Pig
Report reference	PiP16/17
Contributors	Department of Epidemiological Sciences, APHA Epidemiology and Risk Policy Advice team, APHA Information Management Team Data Systems Group, APHA Pig Species Expert Group, APHA
Contacts for queries	erpa@apha.gsi.gov.uk
Data source	eAML2 and ScotEID
Data year	2014-2015

Contents

Who are these reports for?	1
Who did this work?	1
What do the data show about the population?.....	1
How accurate are the data?	1
What do the data not show?	2
How were the maps produced?	3
Annex 1: Data quality statement for pigs (Apr-17)	7
Introduction	7
Overview of source data used.....	7
Overview and purpose of the source data.....	7

Who are these reports for?

These reports are suitable for use in animal health and welfare policy work which requires an estimate of the distribution and size of the pig population at GB level. This type of population level information is often required to assess the economic or social impact of particular animal health policies, for contingency and resource planning, or to provide evidence to trading partners. There are important assumptions and uncertainties with these estimates which the user needs to take into consideration and can be found at Annex 1.

Who did this work?

The Livestock Demographic Data Groups (LDDG) were formed in January 2014 and are made up of APHA representatives from data, epidemiology, species expert and GIS work groups. The LDDGs are grateful to Defra, Welsh Government, Scottish Government, and to AHDB-Pork for their assistance in supplying the eAML2 data and APHA Weybridge DSG staff who handled the ScotEID data through the ScotEID support team.

What do the data show about the population?

The maps (figures 1 and 2) show either the density of animals, with a small map to show how this compares with the density of holdings, or vice versa. The pig and pig holding density maps correlate with the APHA's existing knowledge of the pig population, with expected peaks in pig density in Yorkshire and Humber, the East of England and North-East Scotland, where the majority of large commercial farms exist. The map outputs correspond with high density pig areas on commercial farms identified by the AHDB pig pocketbook 2016 (<http://pork.ahdb.org.uk/media/271528/pig-pocketbook-2016.pdf>, page 6). The maps also highlight a high density of holdings in several other areas of England and Wales, such as the South of England and Wales, where pig density is low, indicating lower numbers of pigs per holding. These areas of low pig density but higher pig holding density may reflect the presence of higher numbers of small holder premises with pigs. This corroborates a similar finding detected in the previous LDDG report, using data from the 2010 Agricultural Census.

How accurate are the data?

Information on the locations of pig holdings and data used to estimate pig density were extracted from a dataset of 24 months of pig movements reported into the electronic animal movement licensing schemes for GB from 2014-2015. Previous use of a 48 month dataset for farm recruitment identified a number of inactive holdings and a 24 month extract was suggested as optimal. The data for England and Wales were downloaded from

the electronic animal movement licensing scheme (eAML2) and sent to APHA by AHDB-Pork who own eAML2. Data for movements between Scottish holdings was accessed from ScotEID (Scottish Electronic IDentification). The schemes record all movements reported by pig owners.

Although the scheme does not record herd size, it was considered to be the most appropriate and accurate data source available due to its ability to identify all holdings to or from which pigs have moved during the 24-month period regardless of size or type of holding. The previous demographics report utilizing the Agricultural Census data which is not targeted to capture data from holdings with few pigs present, showed a total of 10,168 holdings, whereas the pig movement dataset reported 31,663 holdings, indicating the size of the number of holdings missing in the previously used dataset. The LDDG report produced in 2014-2015 highlighted that, in comparison with other suitable datasets, eAML2 consistently matched the highest percentage of holdings in the other datasets and was the only dataset that included most of the British Pig Association holdings, reflecting the greater inclusion of small holdings within the dataset.

The pig movement dataset may contain holdings that no longer have pigs, particularly due to the length (24 months) of the dataset used. This was an issue also relevant for Agricultural Survey data. To evaluate this, an assessment of several time periods (6, 12, and 18 months) was used to detect the size of pig holding population within each extract and determine the types of holdings that were missing when compared to the 24 month dataset. It was expected that holdings with few pigs may be omitted from shorter durations, due to infrequent movements, whereas if large holdings were missing then it would be more likely that these would be missing due to them becoming inactive. The analysis identified that the 24 month dataset was preferential as it maximized the number of holdings recorded, whereas there was minimal change in the number of large holdings that were missing in smaller extracts of time. The 6, 12 and 18 month datasets contained 59%, 74% and 89% of the holdings in the 24 month dataset respectively. The 24 month period selected was a compromise between avoiding inclusion of units that no longer have pigs when a longer time period is used, and ensuring inclusion of units with pigs but infrequent movements which would be missed when a shorter time period is used.

Herd size (and hence pig density) was estimated by applying an algorithm to assess the number of pigs moved from holdings during the 24 month period, which was validated against a subset of accurately matched holdings (2,007) with herd size information held in the 2014 Agricultural Survey. However, inferring herd size from movement data may have introduced inaccuracies. The supporting quality statement provides further detail on the limitations in the data (Annex 1).

What do the data not show?

The representation of the GB pig population by data from eAML2/ ScotEID is recognized to have some limitations. There have previously been errors identified where movements from some breeding herds have not been reported due to a misinterpretation of PRIMO

(Pigs (Records, Identification and Movement) Order) regulations. There is therefore potential for the size of some large breeding herds to be underestimated. The use of a 24 month time period of movements may introduce error, with the number of holdings being an overestimate, as some of these herds may no longer have pigs present.

There is uncertainty inherent in the information displayed. The limitations in the dataset are discussed in the supporting quality statement (Annex 1) and it is important that the users consider these in the context of their work. Population and holding density maps are classified to different scales and units from each other and due care must be taken regarding their interpretation. It should also be noted that the eAML2 data used for this report was extracted by AHDB for this one-off analysis and a regular feed of these data would be preferable, in order to recreate these reports during an outbreak or similar situation.

How were the maps produced?

The maps have been created using the kernel density function in *ArcGIS software*. This tool distributes population information over a defined radius, creating a smooth density surface. Two key parameters that require adjustment are the *search radius distance* and the size of the *output surface grid*. Discussion at the LDDG meetings informed these criteria, and their selection is recognised as a subjective process¹. A search radius of 15km was deemed sufficient to enable distinction between categories and a 1km grid square was used for the density surfaces themselves. The classification bins were limited to six, to aide in cross referencing areas of the map to the key. Note that the ArcGIS Kernel Density tool does not take into account edge effects² and as such density estimates in and around coastal areas may be under estimated.

Comparison between the maps was optimised by assigning similar parameters between the species in this series of reports. However, further refinement of the parameters for each species could represent the information more accurately. Note that the ArcGIS Kernel Density tool does not take into account edge effects, and as such density estimates in and around coastal areas may be under estimated.

In order to produce the maps of pig density, each size category of holding was designated a size weighting value based on cross-reference to a subset of holdings present in the 2014 Agricultural Survey and extrapolated to the full dataset of holdings (Table 1).

¹ Pfeiffer, D. Spatial Analysis in Epidemiology, 2008. p47.

² https://www.e-education.psu.edu/geog586/l5_p15.html

Table 1: Description of the five categories of numbers of pigs moved (either incoming or outgoing movements) related to a holding in a 24 month period, which has been used to estimate relevant herd size categories and to provide weighted values for plotting pig density maps.

Size category of holding	Numbers of pigs moved in 24 month period	Size weighting	Comments
1	1-25	3	Size suggests pet pig owners or small holdings
2	26-300	20	Size suggests small holdings
3	301-2,000	110	Size suggests small commercial farms
4	2,001-8,000	550	Size suggests medium commercial farms
5	8000+	2800	Size suggests large commercial farms

The pig and holding numbers per country were created by assigning the country based on the holding's geographical map reference co-ordinates (Easting and Northing - British National Grid). The spatial coordinates came from the postcode recorded in the dataset. If the holding postcode was missing from the cleaned dataset then the CPH was used to try and identify a holding location from the APHA's operational database known as "Sam". The data shown in table 1 was produced using this method.

Figure 1: Pig population density

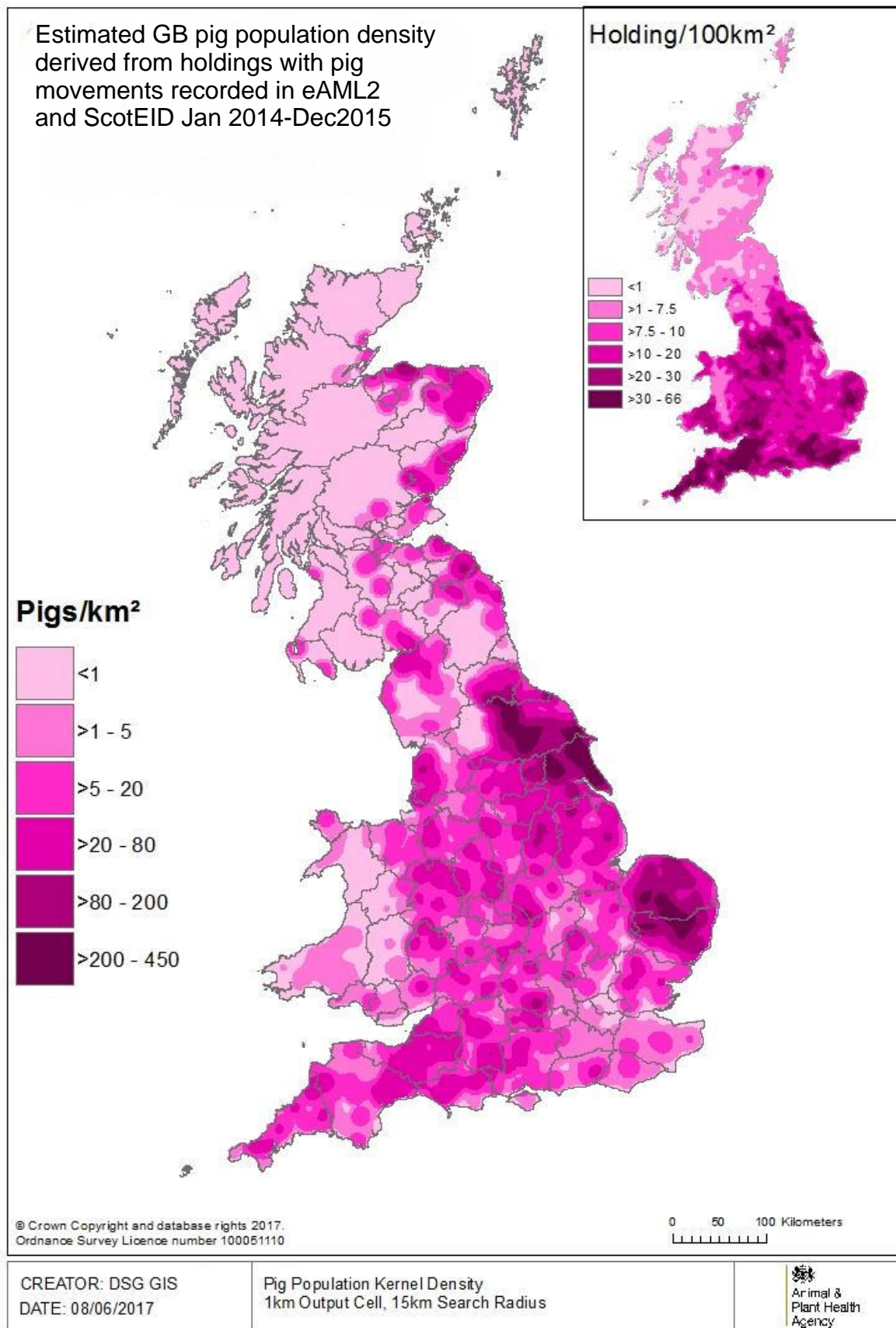
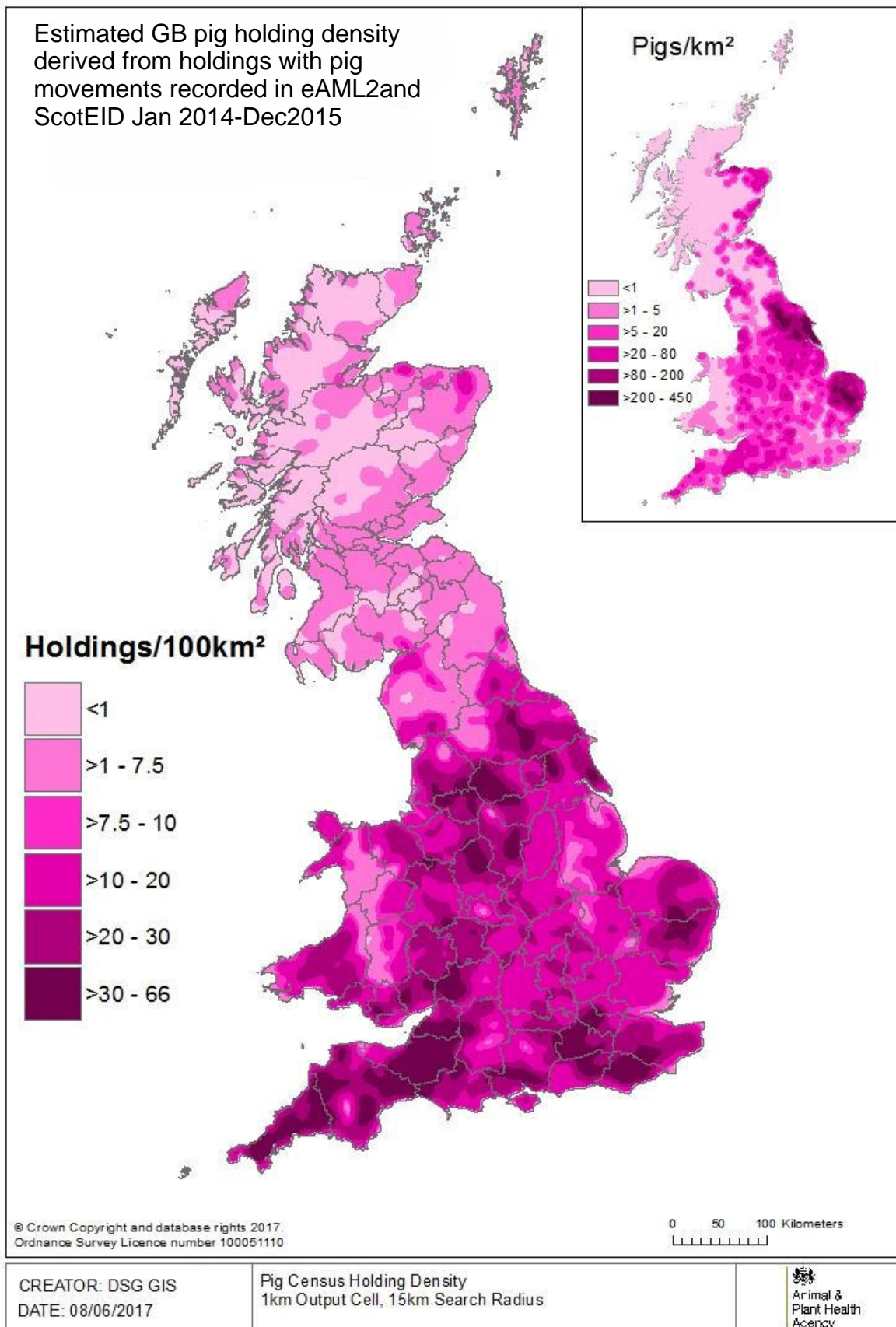


Figure 2: Pig holding density



Annex 1: Data quality statement for pigs (Apr-17)

Introduction

This statement provides an overview of the quality of the data used to underpin the kernel density holding and livestock maps. This statement is written in the context of the data being used to provide an overview of the livestock demographics within Great Britain. The statement may not necessarily relate to data quality for other purposes.

Overview of source data used

Data were supplied by AHDB-Pork from the eAML2 database, accessed through the AHDB PigHub, with movements between Scottish holdings recorded by the ScotEID scheme, supplied by the ScotEID support team.

The eAML2/ ScotEID datasets were chosen to represent the pig data as this has the most inclusive coverage of holdings across GB. The Agricultural Survey was considered, although this survey only holds data on larger agricultural holding and not small holdings.

Overview and purpose of the source data

Both the eAML2 and ScotEID datasets describe the movement of pigs between locations, and records the number and type of pigs moved, which was used to estimated herd size.

Category <i>[definition]</i>	Quality description
Relevance of data <i>[degree to which data meets user needs in terms of currency, geographical coverage, content and detail]</i>	Spatial coverage The data covers GB. Temporal coverage Data were extracted from both sources between March and April 2016 from movements recorded as occurring from January 2014 to December 2015. Key data items available The main data items within the dataset are Date of Movement, Number of Animals Moved and the CPH (county parish holding)

	<p>and Postcode of the departing and destination locations, but other fields are also available within the data.</p>
<p>Timeliness</p> <p><i>[the degree to which data represent reality from the required time point]</i></p>	<p>How often are the data collected?</p> <p>The data are collected continuously throughout the year with users of the two systems registering a movement, which is then finalised after the receiving user confirms receipt of the animals.</p> <p>When do these data become available?</p> <p>The eAML2 and ScotEID databases are live with data continuously being added, although only confirmed movements were extracted, where the receiving holding owner has confirmed receipt of the animals on the database.</p> <p>Data reference period?</p> <p>These data reflect all holding that recorded sending or receiving pigs during 2014 and 2015.</p> <p>How often are the data updated?</p> <p>Once a movement record has been confirmed then these are not changed or updated after import, although holding location details may be updated due to cleaning exercises completed by AHDB.</p>
<p>Accuracy and precision</p> <p><i>[extent of data error and bias and how well data portrays reality]</i></p>	<p>How were the data collected?</p> <p>The data are collected via submissions by registered users via a web portal or a telephone bureau system. Separate movement forms are submitted as movements off and movements on; these are 'paired' by AHDB prior to being made available, i.e. the 'from' and 'to' herd forms are combined into a single record.</p> <p>Sample & collection size</p> <p>There are approximately 30,000 unique CPHs listed in the dataset that had a pig movement.</p> <p>What steps have been taken to minimise processing errors?</p> <p>Data are cleansed by AHDB by comparing holding records with those held elsewhere on the AHDB PigHub. Further cleaning was completed by APHA to remove records with insufficient data to identify a holding and to rationalise holdings that had been</p>

	<p>recorded with varying amounts of identifying information (e.g. movements for a CPH, which had been recorded with and without a postcode, were assigned to the same holding rather than as two separate holdings).</p> <p>What are the non-reporting or non-response rates?</p> <p>We have no information on non-responders, although it has been suggested that some farms within pig breeding companies do not record movements between sites due to a misunderstanding of the requirements.</p> <p>Are any parts of the population unaccounted for in the data collection?</p> <p>It is believed all parts of the population are accounted for. However, commercial breeding farms may have their herd size underrepresented due to the failure to record movements from them to other units within the same pig production company.</p>
<p>Comparability</p> <p><i>[how well these data can be compared with data taken from the same dataset and with similar data from other sources]</i></p>	<p>Within dataset comparability</p> <p>Checks show that data extracted at different times are comparable.</p> <p>Other dataset comparability</p> <p>A comparison of holdings present in eAML2, the Agricultural Survey, APHA's operational database called Sam, the Red Tractor assurance scheme and the British Pig Association (BPA) membership indicated that eAML2 consistently matched the highest percentage of holdings in the other datasets in comparisons and was the only dataset that included most of the BPA holdings (LDDG annual report 2014-2015).</p>
<p>Coherence</p> <p><i>[degree to which data can be or have been merged with other data sources]</i></p>	<p>How consistent is the data over time? If there are differences, what are they and what is their impact? Have there been changes to the underlying data collection?</p> <p>We are not aware of any change in collection methods during recent years but assume minimal bias has been caused.</p> <p>Have any real world events impacted on the data since the previous release?</p> <p>None</p> <p>How have these impacts on the data been managed?</p>

	<p>N/A</p> <p>What other data sources is this dataset comparable with?</p> <p>Other datasets with relevant pig location data available include the Agricultural Survey, Red Tractor and Sam. The Agricultural Survey collects demographical information from a proportion (~30%) of holdings each year, with the remaining population having answers imputed from previous historical records. Holdings included in the Agricultural Survey must meet criteria of a minimum threshold that lists various farmed livestock and crops including criteria of 50 pigs or 10 breeding sows, and so small holdings and hobby farms would be under represented in these data.</p> <p>Red Tractor Quality Assurance scheme is an industry dataset that is regularly updated and includes the number of pigs present on a holding, but only covers commercial herds that use Quality Assured abattoirs and so the holdings present are biased towards large commercial finisher or breeder-finisher farms. Sam is an APHA transactional database which holds a dataset of information regularly collected from farms visited by APHA staff. The information collected would be suitable but there is concern regarding whether the current quality of the data is of sufficient standard, including how up to date are the records and how complete is the coverage of the pig industry. The British Pig Association has a register of pig holdings, although it is believed to be a register of typically smaller pig holdings, such as exotic or rare breeds farms.</p>
<p>Interpretability</p> <p><i>[how well the data is understood and utilised appropriately]</i></p>	<p>Is there a particular context that this data needs to be considered within?</p> <p>This dataset can be used to obtain information regarding animal movements and animal population counts. Although pig numbers fluctuate on farms, it is not believed that a significant seasonal effect would be present that would affect the interpretation of the maps. The dataset was gathered to cover a 24 month period from 2014 to 2015. As registration of movements is legally enforced, we expect the data to be a near complete representation of cattle within the agricultural industry.</p> <p>A “holding” is based on the combination of postcode and CPH (county parish holding) number, which is allocated on a country. However, postcode was not always present and one CPH number may cover multiple sites or post codes. The number of pigs present is based on a categorisation typically based on the number</p>

	<p>of pig moved out of the holding during a two-year period. It should be noted that holdings in this context could be abattoirs, markets and other non-farm locations, and it is assumed that some of the holdings with a size category of 4 or 5 (due to their large number of pigs moved to them) are slaughterhouses.</p> <p>What other information is available to help users better understand this data source?</p> <p>Details of the eAML2 system and a guide on how movements are reported can be found here: https://www.eaml2.org.uk/ami/helpline.eb.</p> <p>Are there any ambiguous or technical terms that may need further explanation?</p> <p>No</p>
<p>Accessibility</p> <p><i>[availability of relevant information and access to the data in a convenient and suitable manner]</i></p>	<p>What data are shared and with whom?</p> <p>Due to restrictions on sharing data provided by a confidentiality agreement between APHA and AHDB, these data cannot be shared without AHDB consent. Where approval for use of data has been provided then data must be aggregated to at least a county level before publishing so individual farms cannot be identified (e.g. by CPH or postcode). Also estimates based on less than five holdings should not be used as this would breach confidentiality.</p> <p>Contact details for data source queries</p> <p>AHDB-Pork: pork.info@ahdb.org.uk</p> <p>ScotEID: help@scoteid.com</p>