

For this bulletin two tests of disclosiveness have been applied. The first test is based on the number of holdings. The second looks for the dominance of the result by one or two large holdings. Any data item that fails either of these tests it is deemed to be disclosive and so suppressed.

At the smaller scales there is clearly a greater risk of statistics being disclosive.

The Welsh Agricultural Survey collects personal information about farmers such as names and addresses. These data are protected by the General Data Protection Regulation, Data Protection Act and the Agricultural Statistics Act. These Acts restrict how the Welsh Government can use this personal information and to whom we can release it. The Acts are, however, concerned with protecting information about individuals. In practical terms they do not restrict the publication of aggregate information where no individuals can be identified.

Data quality issues

The [Code of Practice for Statistics](#) requires the release of data at a level of aggregation that is sufficient to ensure that the results are adequately robust and so “fit for purpose”. This section is designed to be a concise summary *for non-statisticians* of the major data quality issues. A more detailed consideration is attached as [Annex C](#). For more information please [contact us](#).

In this overview there are three key issues to consider:

- Missing values. The Welsh Agricultural Survey is a sample survey. Thus as well as the actual responses we need to estimate results for those farms that were not selected in the sample or that did not respond. These estimated figures mean that a level of uncertainty is attached to the final estimates. Statistical theory allows us to set limits on the general level of uncertainty.
- There is no complete register of farms. This introduces the chance that holdings have been missed. Coverage is best where the holdings are claiming agricultural subsidy. Less robust are specialist poultry, pigs or horticulture producers and very small holdings particularly those run on a non-commercial basis. As with most registers there is also a problem of timing - we might not yet know about some new holdings or for a holding that has ceased to farm we may still believe that it is active.
- Allocation of the farms to small areas. Farms in Wales are not generally compact geographical units. Farms often do not fit neatly into administrative or statistical areas such as local authorities. A farm may have a number of detached fields. Similarly agricultural activity will move around from year to year and season to season. To allow area estimates to be made we have to allocate the activity of a farm to one area. This is usually on the basis of a single “central location”. Clearly this is a great simplification of the real world.

The data for the number of cattle on a farm is taken from an administrative source, rather than survey estimates. The main quality issue for the cattle data is linking farms from the [Cattle Tracing System](#) with the survey data we have about other activities on the farm. This is usually automatic since a common farm reference is used. However, there are a small number of cases where we cannot match a farm from the Cattle Tracing System with one in the survey. The cattle on these farms are included in the published results but not allocated to a Small Area or region.

Note that the June 2010 survey was close to a full census. The survey was sent to around 21,000 farms instead of around 10,000 in a regular year. The increased sample size was to meet EU requirements.

Assessing the impact of the data quality issues

The Welsh Agricultural Survey does not provide a full set of observations for all farm holdings in Wales. It provides estimates at various levels of aggregation. The accuracy of these results depends primarily on the number of actual responses on which the estimates are based. Estimates based on small numbers of observations are inherently less accurate than those based on large sample sizes.

We could in theory eliminate the uncertainty from missing values by ensuring that all farms in Wales completed the survey. The uncertainty from the register could be eliminated by making registration of farms compulsory. This is not to suggest that either of these is either feasible or desirable. Notice however, that this would still leave the uncertainty from allocating the farms to areas.

These data quality issues mean that there is a trade off between the need for local detail and the uncertainty attached to the estimates. This poses a key question for the data user. Does the value to be gained from the apparent extra detail outweigh the additional uncertainty? In practical terms this means making a compromise between the contradictory requirements for precision and accuracy.

No data source can be perfect for all potential uses. For any user an understanding of the data quality is important to give assurance that the data are “fit for purpose”. This may range from a general “ball park” figure up to the need for definitive and comprehensive figures at the smallest level. Data fit for one specific purpose are not necessarily fit for some others.

Recommendations for using agricultural Small Area statistics

It is not possible to produce hard and fast rules on the use of such complicated datasets. However, extensive analysis of the Agricultural Survey suggests that there are seven key observations.

- The Welsh Agricultural Survey provides statistical estimates.
- The Agricultural Survey is a source that is good at giving a broad overview of agriculture in Wales. More detail almost always equates to less accuracy.
- The survey is good at demonstrating trends in the aggregate estimates over time.
- The survey is good at showing the mix of activities on holdings and constructing farm typologies.
- In general the Agricultural Survey produces robust estimates for most of the variables in the survey at *national and regional* levels.
- The nature of the data source means that we can be confident that *on average* the Small Area estimates are reasonably robust over a wide range of variables. This means that it would be reasonable to consider a map of the distribution of a variable across Wales (for example splitting the Small Areas into “high”, “medium” and “low” groups).
- The fact that we have robustness for Small Areas *on average* does not imply that any *specific* Small Area must also be robust. Applications that require accurate estimates for specific localities are likely to be compromised by data quality issues.

Alternative sources of Small Area information

The Agricultural Survey has the advantage of looking at all farming activity and providing consistent series of data. However, it is not the only source of agricultural information. For specific applications, some of the following data sources might be more useful than the survey.

- Direct observation. Never underestimate the power of local knowledge.
- Administrative sources. Including results from agricultural subsidy payments system and other grant schemes. Also includes the existing cattle tracing system and, potentially, the sheep tracing system now being developed.
- Remote sensing. Satellite imagery and aerial photography may help to address land usage questions.
- Sector bodies. Many farming activities have their own organisations. For example farming unions, breed societies, small holdings organisations, horticulture organisations, organic farming registration bodies, levy bodies. These organisations may have data about their specific areas of interest.

Further information

The small area spreadsheet is made available on the [Statistics and Research website](#). Please contact us if you have queries about the data or data quality or if you need the data in some other format.

The spreadsheet is only one way of looking at the data. It has been deliberately designed to give an acceptable level of data quality over a broad range of variables. It is designed with the general user in mind. If users have specific applications in mind and wish to discuss gaining access to alternative Small Area information, please contact the statistician named on the front of the bulletin.

National Statistics status

The [United Kingdom Statistics Authority](#) has designated these statistics as National Statistics, in accordance with the Statistics and Registration Service Act 2007 and signifying compliance with the [Code of Practice for Statistics](#).

National Statistics status means that official statistics meet the highest standards of trustworthiness, quality and public value.

All official statistics should comply with all aspects of the Code of Practice for Statistics. They are awarded National Statistics status following an assessment by the UK Statistics Authority's regulatory arm. The Authority considers whether the statistics meet the highest standards of Code compliance, including the value they add to public decisions and debate.

It is Welsh Government's responsibility to maintain compliance with the standards expected of National Statistics. If we become concerned about whether these statistics are still meeting the appropriate standards, we will discuss any concerns with the Authority promptly. National Statistics status can be removed at any point when the highest standards are not maintained, and reinstated when standards are restored.

Well-being of Future Generations Act (WFG)

The Well-being of Future Generations Act 2015 is about improving the social, economic, environmental and cultural well-being of Wales. The Act puts in place seven well-being goals for Wales. These are for a more equal, prosperous, resilient, healthier and globally responsible Wales, with cohesive communities and a vibrant culture and thriving Welsh language. Under section (10)(1) of the Act, the Welsh Ministers must (a) publish indicators ("national indicators") that must be applied for the purpose of measuring progress towards the achievement of the Well-being goals, and (b) lay a copy of the national indicators before the National Assembly. The 46 national indicators were laid in March 2016.

Information on the indicators, along with narratives for each of the well-being goals and associated technical information is available in the [Well-being of Wales report](#).

Further information on the [Well-being of Future Generations \(Wales\) Act 2015](#).

The statistics included in this release could also provide supporting narrative to the national indicators and be used by public services boards in relation to their local well-being assessments and local well-being plans.

Further details

The document is available at:

<https://gov.wales/statistics-and-research/agricultural-small-area-statistics/?lang=en>

Next update

July 2019

We want your feedback

We welcome any feedback on any aspect of these statistics which can be provided by email to

stats.agric@gov.wales

Open Government Licence

All content is available under the [Open Government Licence v3.0](#), except where otherwise stated.



Annex A

Description of the agricultural small area spreadsheet

Introduction

The spreadsheet contains non-disclosive results for a set of variables from the Welsh Agricultural Surveys from 2002 to 2017. Results are given at the level of the 235 Agricultural Small Areas and the 7 Agricultural Regions.

The spreadsheet is intended to be helpful to users who wish to see the broad distribution of agricultural variables over Wales.

This annex describes what is in the spreadsheet and where it can be found. Further annexes consider the data quality and confidentiality issues (Annex C) and describe the Agricultural Small Areas with a map of the boundaries (Annex B).

Location

The spreadsheet and bulletin can be found on the [Welsh Government website](#).

The link takes you to a page that contains the most recent version of all the agricultural statistics regular outputs, including this bulletin.

Spreadsheet structure

The spreadsheet contains the following worksheets.

- Introduction. A description of the spreadsheet.
- Summary. For each variable shows Wales level statistics for 2002 to 2016: estimated aggregate total; number of holdings; sum of squared values (for calculating the standard deviation); value from actual responses; average value per holding; standard deviation; share of the total estimate derived from actual responses. Table 1 below shows a summary table for the 2016 results.
- Regions. Results for each variable and for each Agricultural Region for each year 2002 to 2017. Note that all these results are non-disclosive.
- SmallAreas. The non-disclosive results for each variable and for each Agricultural Small Area for each year 2002 to 2017. Cells that are disclosive are marked "X".
- Variables. List of the variables used with short notes.
- Lookup. Shows the links between Small Areas, local authorities and Agricultural Regions.

Variables released

The spreadsheet contains a limited range of the variables collected in the June Agricultural Survey. The variables cover the key areas of land usage, numbers of livestock and the numbers of farm workers (not available from 2015 onwards). The spreadsheet also contains counts of the number of farm holdings that have a particular activity. The range of variables has been selected to maximise the amount of non-disclosive information.

The variables in the spreadsheet are summarised in the following table. The table also includes some basic data quality information. The data quality issues are discussed in [Annex C](#).

Geographies used

The results are presented at two geographical scales.

- The 235 “agricultural Small Areas”
- The 7 “agricultural regions”

The derivation of the small areas (including a map of the areas) is described in [Annex B](#). The boundaries for the Small Areas are available on request as a “shape file” that can be read into Geographic Information Systems.

The 7 regions are the ones that have historically been used in the Welsh Agricultural Statistics publication. They are based on groups of local authorities:

- North West – Isle of Anglesey; Gwynedd
- North East – Conwy; Denbighshire; Flintshire; Wrexham
- Powys
- Ceredigion
- Pembrokeshire
- Carmarthenshire
- South Wales – Swansea; Neath Port Talbot; Bridgend; Vale of Glamorgan; Cardiff; Rhondda Cynon Taff; Merthyr Tydfil; Caerphilly; Bleanau Gwent; Torfaen; Monmouthshire; Newport

The tests of disclosure have been applied separately at these two levels. This enables additional non-disclosive information to be released at the regional level.

The spreadsheet also links the Agricultural Small Areas to local authorities. This allows the user to calculate sums for local authorities. A separate calculation of disclosiveness has not been reported for local authority areas. This is largely because of the problems caused by the largely non-agricultural south Wales authorities.

Annex B

Description of the agricultural small areas

The data gained from the June Welsh Agricultural Survey provides the basic physical statistics of farming; areas of land use and of crops; numbers of main livestock types i.e. cattle, sheep and the numbers and kinds of persons working on the farms. All figures relate to the areas of land and numbers of animals and persons on the agricultural holding on the survey date (which is the first weekday in June).

Small areas

Small Area Groups are based on aggregations of Communities within Local Authorities. Where possible, attempts have been made to follow the geographical extent of existing designated boundaries. The 866 Communities within Wales are grouped into 235 Small Areas. A Small Area will contain approximately 100 to 200 farms.

Small Areas were designed to fit into the Unitary Authority boundaries when they were first launched. They were first used with June 1993 results. In order to enable trends in the data to be interpreted the boundaries of the Small Areas have been fixed. Subsequent changes in the Unitary Authority boundaries are therefore not included in the Small Area dataset.

Communities and small area groups

The Welsh Government derives estimates at the individual farm level. However, these cannot be released at this level due to the need to maintain confidentiality and avoid disclosing information about individual farmers.

Previously, information from the agricultural surveys, was published at Country or County/Unitary Authority level. It was also released at District level (in the pre-Unitary Authority structure), however increasing problems in maintaining confidentiality were being experienced as one progressed towards the individual farm holding. In addition, the District was considered too coarse a geographical area to enable effective differentiation.

It was felt that it would be helpful to construct a set of community groups that would be a more detailed level than District but would be large enough to preserve confidentiality. These statistics could then be made available on a regular basis rather than only on request.

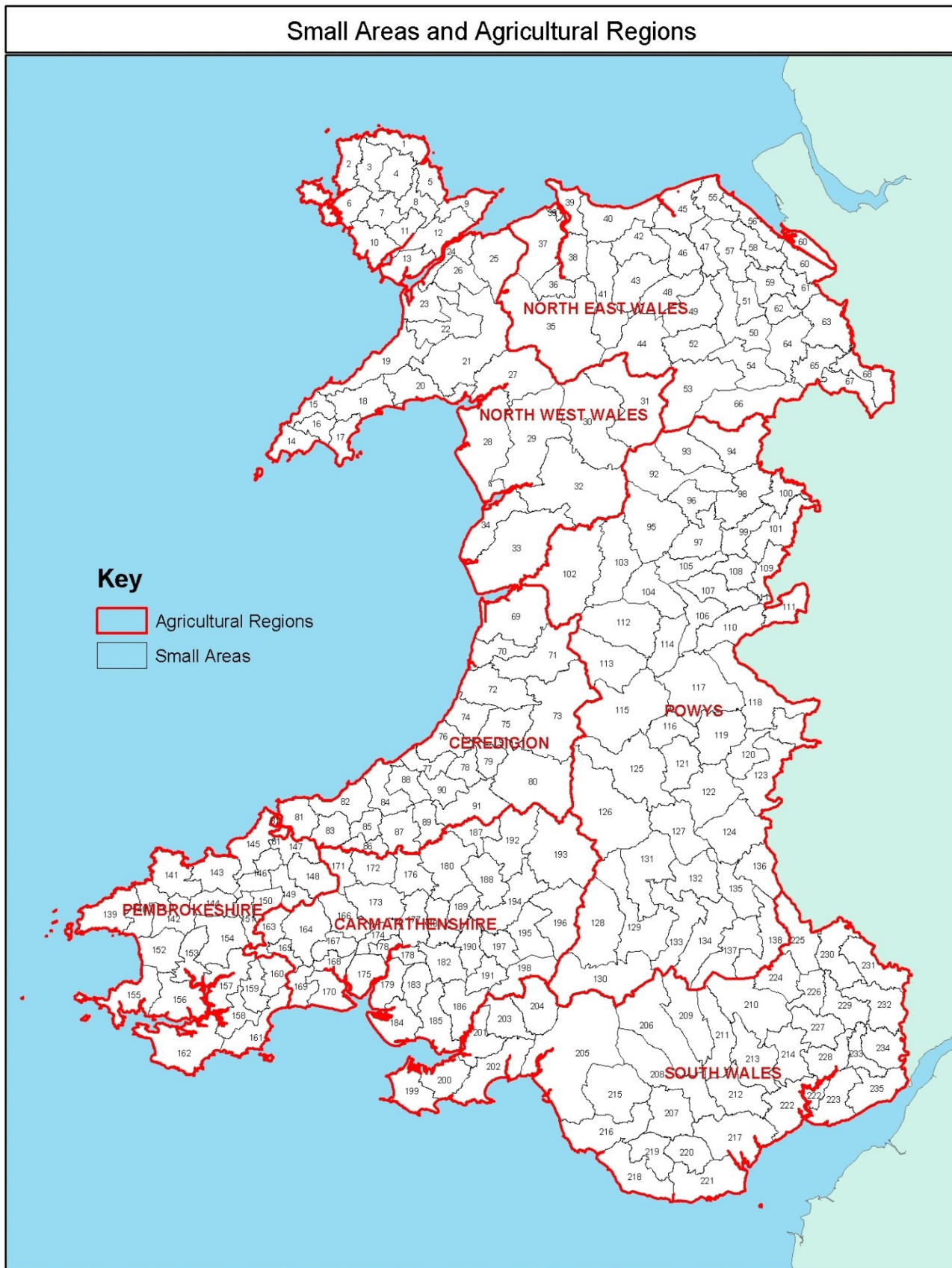
Following local government reorganisation in 1974, Civil Parish Boundaries were reorganised into **Communities** and have differed significantly from the Agricultural Parish Boundaries (**Parishes**). The current Ordnance Survey (OS) administrative boundary data set is based on the Community as the base level of mapping and not the Parish, and it is Communities that are amalgamated to form Unitary Authorities.

Using agricultural Small Areas

[Map 1](#) shows the boundaries of the Welsh agricultural Small Areas, together with the agricultural regions.

The electronic file containing these boundaries ("shape file") can be provided to users who wish to use these data with a Geographical Information System. Please [contact us](#).

Map 1



Annex C

Data quality issues

The following key points should always be remembered.

- The Welsh Agricultural Survey is actually a sample survey not a census.
- The figures thus contain estimates.
- The results are not definitive or comprehensive.
- Where we have an actual response then the values will be generally accurate.
- The accuracy of the estimates for missing data depends very largely on the number of actual responses on which the estimate is based.
- The register of farms on which the survey is based is not a compulsory register and so can only be our best estimate of the true population.
- The survey is based on farms as a building block. These farms are not geographically compact units. We can make estimates for geographical areas but these will contain additional errors because of location issues.

Overview

For those holdings that make a response to the survey we have an accurate record of the agricultural activity. The survey asks for simple land areas, livestock and labour counts. These are variables that any farmer will know – although different terminologies can lead to misunderstandings (particularly in the labour questions).

Since the results are derived from a sample survey they are subject to various sources of uncertainty. The key sources of uncertainty are:

- Missing data.
- Register of farms.
- Geographical location of farms.

Missing data

The June Agricultural Survey is a sample survey of farms. Not all farms are selected for the sample. Not all farms that are selected complete the questionnaire. Standard statistical techniques are used to estimate the Wales total figures, including the missing holdings.

[Table 1](#) in Annex A shows the proportions of the agricultural variables in the released data set that come from actual responses. The actual value of the share from responses changes from variable to variable and between surveys. As a rule of thumb, the share usually falls between 40% and 50%. *The remainder will be imputed and so subject to a degree of sampling error.*

The overall Wales totals are estimated from the survey results by assuming that the increase from last year for the population as a whole is the same as that we see in the sample. Given that we have taken a

random sample of farms this assumption is supported by statistical theory. Analysis of actual results also shows that *at the aggregate level* these estimates are adequately robust.

These estimates will be at their best when there are a lot of similarly sized farms that are, broadly speaking, following the same patterns. This is usually the case for cattle and sheep farms in Wales. The estimates are more problematic when an activity is dominated by a few large farms or when there is great volatility. Specialist chicken producers would be an example of the former while very small or non-commercial farms would be an example of the latter.

To enable a full set of results to be produced – including cross tabulations of one variable against another – we attempt to impute missing values for farms that were not in the survey or did not provide a response. These imputed farm values are constructed so that they equal the Wales total estimates. The imputations are made in the same way as the total estimates by assuming that the whole population and the sample have the same increase over last year.

A limitation of this methodology is the assumption that an imputed farm will have the same set of agricultural activities as it did in the base year. For the individual holding there can be no changes to the types of crops grown or livestock kept – although the areas and numbers will change.

Clearly these imputed farm level results are not particularly accurate *for individual farms*. Their usefulness comes in being a flexible building block to construct a very wide range of aggregate estimates. The accuracy of these aggregate estimates comes from combining a large number of individual farms. Statistical theory shows how this process of aggregations leads to more accurate estimates.

Register of farms

There is no compulsory register of farms in the UK. The voluntary nature of the register means that it will always have issues about missing farms (false exclusions) and farms that have ceased to operate but that are still on the register (false inclusions). If nothing else there will be issues of how quickly we are informed of changes.

There are three main routine sources of information for the register.

- Links with administrative systems. The register is updated with information from various administrative systems. The main sources are the subsidy payment system and the compulsory register of cattle keepers (BCMS). Data are received from time to time from other administrative systems.
- Correspondence through running the various agricultural surveys. This provides updated name and address details plus a number of forms that are returned undelivered.

- The agricultural magazine that the Welsh Government publishes known as “[Gwlad](#)” (the Welsh for “countryside”). The statistical register is used as the basis for the magazine mailing lists. Farmers notify the magazine of name and address changes. They can also ask to be removed from the circulation list. This is important for the register because farmers see the magazine as a benefit that they want to continue to receive. Improving the quality of the statistical register is a welcome bonus from this fact.

From these sources it is clear that register information will be at its best when we have regular contact with the farms. The coverage will be at its best for farms that are within the agricultural subsidy system because they have the strongest reason to be registered. In Wales the majority of farms are claimants.

For holdings outside the subsidy system there are four key types of holding that are known to have coverage problems.

- Specialist poultry producers.
- Specialist pig producers.
- Specialist horticulturists.
- The very smallest holdings including non-commercial concerns.

The specialist producers are a particular concern because there tend to be few of these in Wales and these sectors are dominated by a small number of farms. The most extreme is for chicken producers where more than 90% of the birds are on less than 1% of the farms with chickens.

The smallest farms do not make a significant contribution to most agricultural variables. Typically these holdings will have well below 1% of the Wales total. Note that the very smallest farms contribute a significant proportion of the total number of horses in Wales.

There are, however, large numbers of these very small farms. This may be important if what we want to know is all holdings with a particular activity – for example livestock keepers at the time of the Foot and Mouth outbreak.

Geographical location of farms

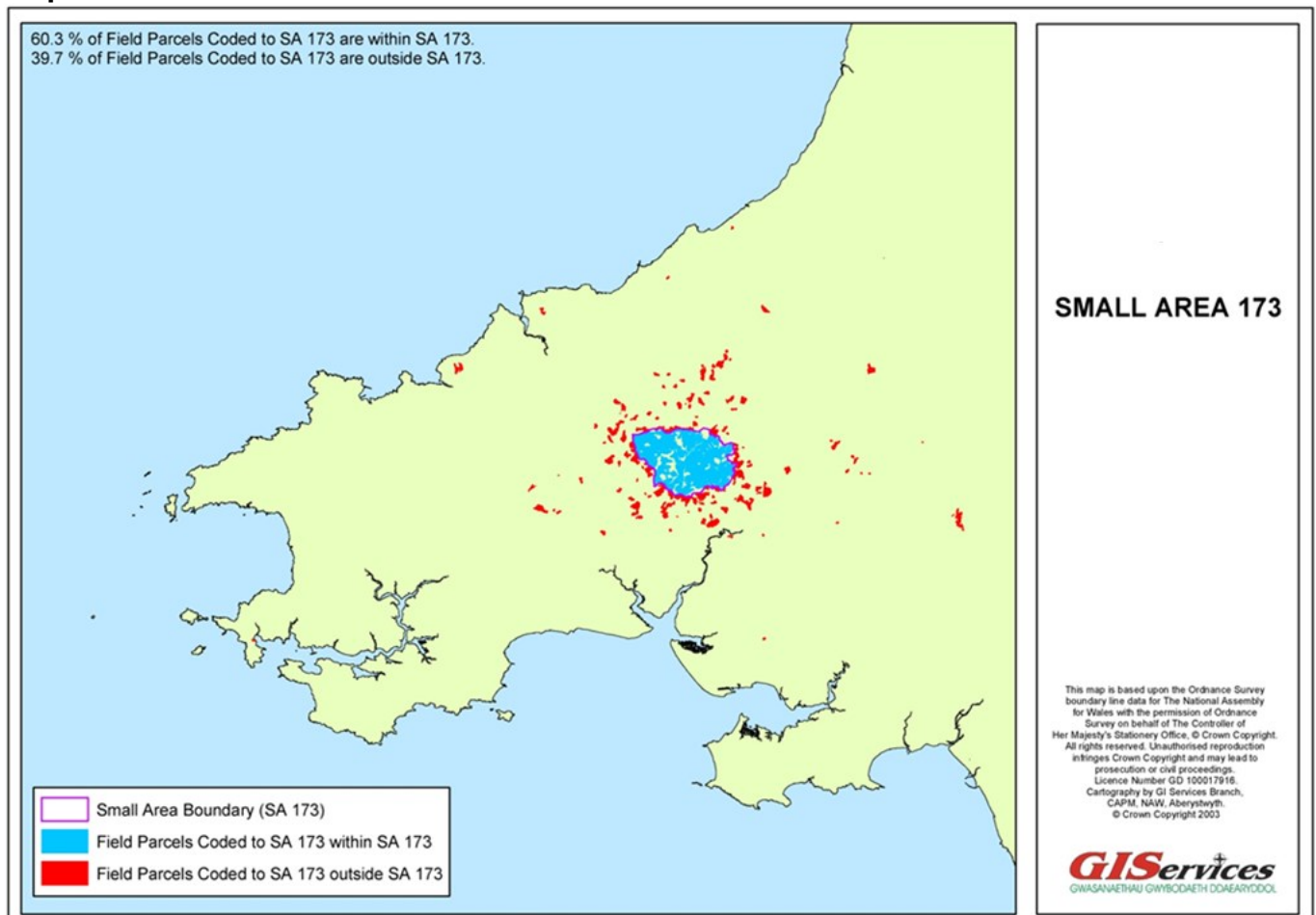
The Welsh Agricultural Survey collects information at the level of the individual farm. These farms are collections of land farmed together as a unit. *They are not necessarily contiguous sets of fields.* The fields on a farm may be scattered over a wide area. Therefore, while we will know the number of animals on a farm we do not know where the animals graze. Similarly while we know the total area of crops we do not know where the crops are grown.

It is possible to make an estimate of the general location of a farm. These estimates can use the digital field boundaries where these are available through the subsidy system and address information otherwise. We can then build up geographical estimates by assuming that all the activity on a farm takes place at this estimated location. *Clearly this an estimate and a great simplification of reality.*

The accuracy of the geographical estimates depends mainly on using areas that contain a sufficiently large number of farms. With a large number of farms the overall net error will decrease as a false inclusion on one farm is balanced by a false exclusion on another. However, for very specific location, with small numbers of farms involved, the level of error because of the location issues must be more significant.

The significance of this problem can be seen in [Map 2](#). This uses the individual field registrations from the subsidy payments system. The map shows all the registered fields that belong to farms that are thought to fall within the single Small Area.

Map 2



Source: Welsh Government

