

Quality assurance of administrative data report for the agriculture industry, September 2017

Investigation of the administrative data sources used in the production of short-term economic output indicators by ONS's National Accounts and Economic Statistics Group.

Contact:
Tanzila Rahman
stoi.development@ons.gsi.gov.
uk

Release date:
4 September 2017

Next release:
To be announced

Table of contents

1. [Introduction](#)
2. [Quality assurance assessment](#)
3. [Areas of quality assurance](#)
4. [Summary](#)

1 . Introduction

1.1 Background

National Accounts and Economic Statistics (NAES) Group within Office for National Statistics (ONS), request data from the Department for Environment, Food and Rural Affairs (Defra) on the agriculture industry. This includes two activities, namely the production of crop products and production of animal products. These data form one source in the calculation of short-term economic output indicators namely gross domestic product (GDP (O)).

This report outlines the process data take from initial collection through to the output of the release. It identifies potential risks in data quality and accuracy as well as details of how those risks are mitigated.

This report forms the latest in a series of quality assurance reports produced by the NAES Group to interrogate the data sources we use. As such this report specifically focuses on the agriculture industry and does not aim to cover other industries that make up the short-term economic output indicators, many of which will be considered in due course.

Further information relating to quality and methodology for the short-term economic output indicators can be found in our GDP [preliminary estimate release](#) and [UK Index of Services Quality and Methodology Information](#).

1.2 Standard industrial classification (SIC) overview

The agriculture industry covers the activities under [UK Standard Industry Classification \(SIC\) 2007](#) division 01. This division includes crop and animal production, hunting and related services.

Based on the UK SIC 2007 the agriculture industry is classified into six groups:

01.1 Growing of non-perennial crops

- 01.11 - Growing of cereals (except rice), leguminous crops and oil seed: barley, wheat, potatoes, and oilseed rape
- 01.12 - Growing of rice
- 01.13 - Growing of vegetables and melons, roots and tubers: horticulture
- 01.14 - Growing of sugar cane
- 01.15 - Growing of tobacco
- 01.16 - Growing of fibre crops
- 01.19 - Growing of non-perennial crops

01.2 Growing of perennial crops

- 01.21 - Growing of grapes
- 01.22 - Growing of tropical and subtropical fruits
- 01.23 - Growing of citrus fruits
- 01.24 - Growing of pome fruits and citrus fruits
- 01.25 - Growing of other tree and bush fruits and nuts
- 01.26 - Growing of oleaginous fruits
- 01.27 - Growing of beverage crops
- 01.28 - Growing of spices, aromatic, drug and pharmaceutical crops
- 01.29 - Growing of other perennial crops

01.3- Plant propagation

- 01.30 - Plant propagation

01.4 Animal production

- 01.41 - Raising of dairy cattle: cattle, milk
- 01.42 - Raising of other cattle and buffaloes
- 01.43 - Raising of horses and other equines
- 01.44 - Raising of camels and camelids
- 01.45 - Raising of sheep and goats: sheep
- 01.46 - Raising of swine/pigs: pigs
- 01.47 - Raising of poultry: poultry and eggs
- 01.49 - Raising of other animals

01.5 Mixed farming

- 01.50 - Mixed farming

01.6 Support activities to agriculture and post-harvest crop activities

- 01.61 - Support activities for crop production
- 01.63 - Support activities for animal production
- 01.62/1 - Farm animal boarding and care
- 01.62/9 - Support activities for animal production (other than farm animal boarding and care)not elsewhere classified
- 01.63 - Post-harvest crop activities
- 01.64 - Seed processing for propagation

01.7 Hunting, trapping and related services activities

- Hunting, trapping and related services activities

According to the [Inter-Department Business Register](#) (IDBR)¹ there were 139,605 enterprises classified under division 01 in 2016. This was a slight increase of approximately 530 enterprises (0.38%) from the previous year (2015).

SIC 01.45 raising of sheep and goats and SIC 01.50 mixed farming made up the majority of the industry with 28,635 and 29,345 enterprises in each class respectively (20.51% and 21.02% of whole industry).

The annual turnover of 48,640 (38.4%) enterprises within division 01 was less than £50,000. The vast majority of enterprises within division 01; 125,045 (89.6%) had registered turnover of below £1 million.

Within division 01, there were 123,450 enterprises with fewer than five employees (88.42% of the division).

Notes for : Introduction

1. The Inter-Departmental Business Register (IDBR) is a comprehensive list of UK businesses that is used by government for statistical purposes. It provides the main sampling frame for business surveys carried out by both ONS and other government departments. It is also a key data source for analyses of business activity.

2 . Quality assurance assessment

2.1 UK Statistics Authority Quality Assurance Toolkit

The assessment of this data sources has been carried out in accordance with the [UK Statistics Authority Quality Assurance of Administrative Data \(QAAD\) Toolkit](#).

Each data source is evaluated according to the toolkit's risk and profile matrix reflecting the level of risk of data quality concerns and the public interest profile of the statistics. See Table 1.

Table 1: UK Statistics Authority administrative data quality assurance risk and profile matrix

Level of risk of Quality concerns	Public interest profile		
	Lower	Medium	Higher
Low	Statistics of lower quality concern and lower public interest [A1]	Statistics of low quality concern and medium public interest [A1/A2]	Statistics of a low quality concern and higher public interest [A1/A2]
Medium	Statistics of medium quality concern and lower public interest [A1/A2]	Statistics of medium quality concern and medium public interest [A2]	Statistics of medium quality concern and higher public interest [A2/A3]
High	Statistics of higher quality concern and lower public interest [A1/A2/A3]	Statistics of higher quality concern and medium public interest [A3]	Statistics of higher quality concern and higher public interest [A3]

Source: Office for National Statistics

The toolkit outlines four specific areas for assurance and the rest of this report will focus on these areas. These are:

- operational context and administrative data collection
- communication with data supply partners
- quality assurance principles, standards and checks applied by data suppliers
- producer’s quality assurance investigations and documentation

In the assurance of our data source, we have chosen to give a separate risk and profile matrix score (Table 2) for each of the four areas of assurance, this will allow us to focus our investigatory efforts on areas of particular risk or interest to our users. Scoring for the agriculture industry is shown in Table 2.

2.2 Assessment and justification against the QAAD risk and profile

Table 2: Quality assurance risk and profile matrix assessment used to measure agriculture

	Low	Medium	High
Operational context and administrative data collection		A2	
Communication with data supply partners		A2	
Quality assurance principles, standards and checks by data supplier		A2	
Producers quality assurance investigations and documentation	A1		

Source: Office for National Statistics

Quality concerns and level of public interest have both been set as “Medium” due to the contribution that the agriculture feeds into gross domestic output (0.67%) and the increase in the public profile of agriculture since the UK voted to exit from the European Union. As such, a score of A2 is deemed appropriate for this data source.

All scoring was carried out by National Accounts and Economic Statistics (NAES) Group based on the level of risk of the data and interest of our users. Results for each area of assurance for agriculture are shown in Table 2. If you feel that this report does not adequately provide this assurance then please contact stoi.development@ons.gsi.gov.uk with your concerns.

3 . Areas of quality assurance

3.1 Operational context and data collection – matrix score, A2

This relates to the need for statistical producers to gain an understanding of the environment and processes in which the administrative data are being compiled and the factors that might increase the risks to the quality of the administrative data.

3.1.1 Overview of the components

Department for Environment, Food and Rural Affairs (Defra) are responsible for collecting National Statistics data on agriculture. There are a variety of data that Defra collects, however, this report will only focus on data that National Accounts Economic Statistics (NAES) Group uses for gross domestic product (GDP) outputs.

Defra supply 11 data components (Table 3) to NAES, five are on a monthly basis, one is supplied quarterly and the remaining five are provided annually. All components collected by Defra, which are used by NAES, with the exception of horticulture, have [National Statistics](#) status and therefore conform to the statistical standards outlined by the [Code of Practice for Official Statistics](#).

Table 3: Component by data type, reporting period and availability to NAES

	Component	Data type	Reporting period	Availability
1	Wheat	Volume (thousand tonnes)	Annually	October (forecast) / April (final)
2	Barley	Volume (thousand tonnes)	Annually	October (forecast) / April (final)
3	Oilseed rape	Volume (thousand tonnes)	Annually	October (forecast) / April (final)
4	Potatoes	Volume (thousand tonnes)	Annually	October (forecast) / April (final)
5	Cattle	Volume (thousand tonnes)	Monthly	6 weeks after the end of the month
6	Sheep	Volume (thousand tonnes)	Monthly	6 weeks after the end of the month)
7	Pigs	Volume (thousand tonnes)	Monthly	6 weeks after the end of the month
8	Poultry	Volume (thousand tonnes)	Monthly	6 weeks after the end of the month
9	Eggs	Volume (thousand tonnes)	Quarterly	Feb / May Aug / Nov
10	Milk	Volume (thousand tonnes)	Monthly	6 weeks after the end of month
11	Horticulture and other crops*	Constant Price	Annually	October (forecast) / April (final)

Source: Office for National Statistics

Notes:

Other crops include (oats, and summer cereal mixtures, rye, other cereals, plants and flowers, fruit, oilseed rape, other oilseed, sugar beet, fibre plants, Vegetables, Hops, Other industrial crops, forage plants, other crops including seed, protein crops).

3.1.2 Data sources

Other than horticulture all of the 11 components mentioned in Table 3 are collected by surveys. Table 4 gives an overview of the seven data sources used by the relevant component.

Table 4: DEFRA data sources used to collect each component

Component	Data source(s)	Methods
1 Wheat, barley and oilseed rape	<p>'Cereal and Oilseed Rape Production survey' for England</p> <p>'Cereal production and disposal survey' for Scotland</p> <p>Welsh production is estimated on a regional basis within Wales from the 'June Survey of Agriculture and Horticulture' along with yields for the English regions bordering Wales</p> <p>Northern Ireland production figures are based on the 'June Survey of Agriculture and Horticulture' and are derived from cereal yields (Post-Harvest survey) and then combined with the respective cereals areas from the latest census</p>	Surveys
2 Potatoes	Family Food Survey, Great Britain and Northern Ireland	Surveys
3 Cattle, sheep and pigs	<p>Monthly Slaughterhouse Survey, England and Wales</p> <p>'June Agricultural Census', 'Rural and Environment Science and Analytical Service (RESAS), Scotland</p> <p>'Northern Ireland Agriculture survey', Department of Agriculture, Environment and Rural Affairs (DAERA), Northern Ireland</p>	Surveys
4 Poultry	<p>Survey of England and Wales Poultry Slaughterhouses</p> <p>'June Agricultural Census,' in RESAS, Scotland</p> <p>Northern Ireland poultry statistics</p>	Surveys
5 Eggs	<p>'Defra Hatchery Survey layer chick placings', UK</p> <p>Quarterly survey of registered UK egg packaging statistics called 'Egg Packers survey', UK</p>	Surveys
6 Milk	<p>Monthly and quarterly survey of milk availability and usage by dairies in England and Wales</p> <p>Monthly survey for milk production, RESAS for Scotland ³</p> <p>Monthly survey for raw milk production, DAERA for Northern Ireland⁴</p>	Surveys
7 Horticulture	<p>Horticulture market reports of wholesale prices from four leading wholesale markets in England and Wales</p> <p>'June Agricultural Census' and 'December Agriculture survey, RESAS Scotland</p> <p>Statistical Review of Northern Ireland Agriculture Northern Ireland</p>	Surveys

Source: Office for National Statistics

The links for the latest methodology papers for each of the components can be found in Table 5.

Table 5: Methodology papers for the components

Component	Web I nks
1 Wheat, barley and oilseed rape	Farming Statistics 2016
2 Potatoes	Family Food 2015
3 Cattle, sheep and pigs	United Kingdom Slaughter Statistics
4 Poultry	United Kingdom Poultry Meat Statistics
5 Eggs	United Kingdom Egg Statistics
6 Milk	Usage of milk by dairies in England and Wales
7 Horticulture	June Survey of Agriculture and Horticulture: Methodology

As the survey process by which all data for the different components is collected share commonality, we have decided to focus on the cattle, sheep and pigs component as a specific case study to describe the wider process of data collection for all components. Where data collection of other components does differ will be considered in section 3.1.3.4. Please note, as we are aware that the process of data collection for potatoes and horticulture differs throughout we will be considering these components separately in their own sections.

We believe this approach is the most appropriate to give our users specific details of the data's journey whilst improving clarity by reducing repetition.

3.1.3 Cattle, sheep and pigs

The production data for cattle, sheep and pigs are collected monthly from slaughterhouses (abattoirs) across the UK and are collected by separate surveys for each country. The data for England and Wales is collected from abattoirs by Defra and is then supplemented by data from the Food Standards Agency (FSA). The Scotland survey is collected by RERAS (Rural and Environment Science and Analytical Service) and the Northern Ireland data is collected by DAERA (Department of Agriculture, Environment and Rural Affairs). RESAS and DAERA provide the data to Defra. Defra then provide the UK cattle, sheep and pigs production figures to NAES, which is used in the monthly GDP.

A process map showing the collection of cattle, sheep and pigs data' visually is shown in Figure 1 of the annex.

3.1.3.1 Methodology

The production figures for cattle, sheep and pigs in England and Wales is collected by a statutory monthly survey, run by Defra, of registered slaughterhouses, under Regulation number EC.1165/2008. Detailed information on this legislation is available by selecting [Animal production](#) on the Eurostat website.

Information on livestock slaughter numbers and weight of meat produced are collected from approximately 91 slaughterhouses. All the major slaughterhouses⁵ participate in this survey and because the survey is statutory there is usually a 100% response rate.

This monthly slaughterhouse survey accounts for 90% of the slaughterhouse throughput for each livestock type. Smaller slaughterhouses are then supplemented by administrative data from the Food Standards Agency (FSA), which provides a complete monthly coverage of livestock slaughtering⁶. This means that Defra have 100% coverage of the industry.

In order to calculate production Defra use the information on the weight of meat produced from a sample of animals weighed at slaughterhouses. The survey respondent consists of around 55 slaughterhouses. This information is then used to provide average dressed carcass weights for each animal type and is calculated by adding up the total weight of the meat produced and dividing by the number of animals weighed.

The volume of meat production is estimated from the number of all livestock slaughtered and average dressed carcass weight information collected. Livestock imported into the UK for slaughter is also included in the total "Home killed" production.

As agriculture and its practices are devolved, similar surveys are also run by RESAS in Scotland and by DAERA in Northern Ireland and used by Defra to collect UK data as a whole. Scottish statistics on livestock slaughtering are available in the [Economic Report on Scottish Agriculture](#). and Northern Ireland results for [cattle and sheep](#), and [pigs](#) are available on the DAERA website.

Defra publish the statistics for cattle, sheep and pigs in their statistics bulletin within 3 weeks after the end of each month (see link in Table 5). NAES receives the data via a spreadsheet from Defra monthly, 6 weeks after the end of the month.

3.1.3.2 Revisions

The tight deadlines to publish mean that the statistics that Defra produce are subject to revisions. Information on any revisions to previously published information is provided in Defra's statistics notice and associated datasets.

Defra also provide briefings to NAES accompanying the data if there are fluctuations based on historical trends to explain these.

The revisions to the cattle, sheep and pigs statistics could occur for various reasons, including:

- if survey data has not been received by Defra's deadlines then estimates are made based on respondents' previous monthly returns; the estimates are replaced when actual survey data is received
- survey respondents sometimes supply amended figures for previous periods
- if administrative data for the smaller slaughterhouses have not been received then Defra will make an estimate from the previous data received; this is then replaced when the actual data is received

Revisions for wheat, barley, oilseed rape, poultry, eggs and milk could occur for the same reasons mentioned previously.

3.1.3.3 Operational context of cattle, sheep and pigs data source

Defra have a long-standing history of using surveys to collect this data. The advantage of this is that it enables Defra to improve accuracy by making comparisons against historical trends. Defra still believe these surveys are the best source to collect data on our components; however, they acknowledge that the use of surveys has to provide timeliness, accuracy and value for money and have therefore evaluated various other data sources such as the possibility of using satellite mapping, which is still being monitored. This shows that Defra are engaged in the data collection process and willing to improve and innovate.

3.1.3.4 Operational context of other data sources

As mentioned previously the process by which the data is collected for cattle, sheep and pigs is also consistent for six other components: wheat, barley, oilseed rape, poultry, eggs and milk.

The components wheat, barley, oilseed rape, poultry, eggs and milk are all collected by similar surveys. Defra confirmed that the surveys are a representative sample providing 90% coverage of these components. They believe the coverage is sufficient to minimise bias as the sample used focuses on the larger players within the industry. The remaining 10% tend to be the small players in the industry and do not make a significant contribution to agriculture.

The statistics for milk, poultry, wheat, barley and oilseed rape are collected by statutory surveys:

- the statistics for milk are required monthly under Council Directive 96/16EC
- the statistics for poultry are collected under Regulation EC 1165/2008 (Animal production)
- the statistics for wheat, barley and oilseed rape are also published in [Agriculture in the United Kingdom \(AUK\)](#), which fulfils the requirement under the Agriculture Act 19939

The survey for egg production volumes are not required by law but are still collected as National Statistics.

3.1.4 Potatoes

The method for collecting production of potatoes has recently changed. The production figures for potatoes are now modelled on the results from Defra's [Family Food](#) publication. The latest statistics for potatoes come from the 'Family Food 2015' report, which is based on the 2015 family module of the Living Costs and Food Survey (LCFS) which is run by Office for National Statistics (ONS).

The LCFS sample for Great Britain is a multi-stage stratified random sample with clustering. The sample is obtained from the Small Users file of the Postcode Address File – which is the Post Office's list of addresses. The Northern Ireland sample is a random sample of addresses from the Land and Property Agency list. The survey is a voluntary survey of private households. For Great Britain 11,484 addresses were selected for LCFS and 10,349 were eligible¹⁰ – of these 4,760 cooperated fully in the survey and the overall response rate was 46%. In Northern Ireland 158 households co-operated out a sample of 255 with a response rate of 62%. The response rates are in line with other major government surveys.

[Further information about the survey methodology](#) and the [results for this report](#)¹¹ are available.

In this case the change in methodology was not effectively communicated with NAES. This is being addressed through enhanced communication methods (discussed further in section 3.2).

3.1.5 Horticulture

The methodology by which horticulture price data is collected differs by country within the UK. The data for England and Wales is collected directly by Defra, for Scotland the horticulture data is collected by RESAS and for Northern Ireland the data is collected by main industry contacts.¹³ Defra then provide a full dataset of UK horticulture data to NAES.

The next section will further explore these regional differences.

A process map showing the collection of horticulture data visually is shown in Figure 2 of the annex.

3.1.5.1 Methodology

For England and Wales horticulture data is collected by Defra and obtained from price data for a range of fruit and vegetables from wholesale markets. Price data is the national average of the most usual prices charged by wholesalers for selected home-grown fruit and vegetables. Four leading wholesale markets in England are selected, that is, Birmingham, Bristol, Liverpool and New Spitalfield (in London) and collected weekly (prices for Wales are assumed to be the same as England). Data is collected by market reporters visiting these wholesale markets each week, reports are then emailed to Defra. Statistics are published weekly in the [National Average Wholesale Prices of Selected Home-Grown Horticultural Price](#).

[More information about the methodology](#) is available.

Each wholesale fruit and vegetables price is adjusted by a factor that takes into account the relative proportions and values of produce going through various supply chains to the market, that is, this includes supermarkets, processing, direct sales on-farm as well as through wholesale markets. This price is then adjusted to take into account commission and handling and carriage costs to then derive an estimate of the farm-gate price.¹⁵

The factors mentioned previously and the need to keep up-to-date information is kept under review by Defra. However, analysis has shown that the final valuation estimates are not highly sensitive to these factors. This is true especially considering the wider limitations in using wholesale market price data (see section 3.1.5.3).

Scotland

The horticulture price data for Scotland is collected by the June census and December survey of farms. Prices are estimated using the previous year's survey data adjusted by the latest price data from Glasgow Market price data. Price data for strawberries and blackcurrants are derived from a postal survey of horticultural units. Further information on Scottish horticulture is available in the [reports](#) and [methodology documents](#).

Northern Ireland

Horticulture is a relatively small contributor to the value of agricultural output in Northern Ireland contributing around 3 to 4% of the total value of output. Farm-gate price data are obtained through main industry contacts. These are then cross-compared across sources and against general trends and take into account wider agronomic and market knowledge of each of the different sectors to ensure the data are credible and representative.

Price data take into account the end-use of the produce and any price differentials in order to derive a representative weighted annual average price. The volume of output takes into account marketings in each particular year. For example, for crops such as apples where the crop is marketed over 2 years, it will take into account marketings from two seasons for each particular calendar year.

The [Statistical Review of Northern Ireland Agriculture](#) has detailed information on prices and value of production for fruit (primarily apples), vegetables, mushrooms and ornamental and hardy nursery stock.

3.1.5.2 Revisions

Revisions could occur for various reasons, including when data from third parties is unavailable or provisional at the time of publishing. Changes can also occur if new data becomes available or if there are discrepancies found in the datasets. When Defra send ONS the new data it will also includes any revisions to the annual data.

3.1.5.3 Operational context of the horticulture source

Defra accept that there are some limitations of using valuation estimates of horticulture that they provide to NAES. Defra stated that these estimates should be treated with caution as they are derived from wholesale market prices with several factors applied to derive an approximate farm-gate price. These estimates therefore will not capture the full range or complexity of pricing and marketing arrangements but should still provide a reasonable estimate of broad trends.

Despite these limitations, wholesale market price data provides the best available route for deriving farm-gate price or valuation estimates as it is widely available and cost-effective to collect. This is especially true given commercial sensitivities concerning the sourcing of reliable and up-to-date price data.

Strengths

- there is 100% coverage for cattle, sheep and pigs
- the data provide 90% coverage for the components wheat, barley, oilseed rape, poultry, eggs and milk
- the data for cattle, sheep and pigs, poultry and milk is collected by statutory survey
- statistical releases produced from these collections have National Statistics status, ensuring compliance with the Code of Practice
- the long-standing use of surveys allows Defra to make historical comparisons
- the data for horticulture is widely available and is cost-effective to collect
- the statistics have a wide number of users and Defra work closely with their users to optimise collection.
- Defra inform NAES with briefings on the data submitted when required

Weaknesses

- in order for the data to be relevant to users Defra publish statistics to a tight deadline, resulting in revisions
- agriculture and its practice are devolved so data are collected by different surveys in different countries, leading to slight differences in methodologies and more opportunities for errors to be introduced
- horticulture price data must be treated with caution as captures approximate farm-gate prices and not the full range or complexity of pricing and marketing arrangements
- change in methodology around the collection of potato data not effectively communicated

Next steps

- maintain this level of knowledge on the subject area
- work with Defra to explore the potential for improvement to horticulture estimates

3.2 Communication with data supply partners – matrix score, A2

This relates to the need to maintain effective relationships with suppliers (through written agreements such as service level agreements or memoranda of understanding), which include change management processes and the consideration of statistical needs when changes are being made to relevant administrative systems.

A process map showing the flow of communication between data supplier to output producers' is shown visually in Figure 3 of the annex.

3.2.1 Communication between Department for Environment, Food and Rural Affairs (Defra) and their data suppliers

Defra (Department for Environment, Food and Rural Affairs), Rural and Environmental Science and Analytical Service (RESAS) and Department of Agriculture, Environment and Rural Affairs (DAERA) send out forms electronically via email and this is accompanied by information that explains the purpose of the survey and provides guidance to aid completion. There is also a helpline facility, although this is generally not used as respondents are familiar with the forms.

For the components wheat, barley, oilseed rape, poultry, eggs and milk, non-respondents are chased up by email and then by phone. For cattle, sheep and pigs, if they do not respond then Department for Environment, Food and Rural Affairs (Defra) use the Food Standards Agency data.

Defra communicate with their data suppliers when quality-assuring the data they receive from the survey results. If Defra have any concerns about the validity of the results in comparisons with historical results they will contact the individual supplier by phone or email and query the result.

Due to the limitations of the horticulture data, Defra have annual face-to-face meetings with RSK ADAS, who provide the horticulture data, in order to address any ongoing issues with the data. Furthermore the prices team within Defra will contact the markets over any queries on the data within a day of receipt.

3.2.2 Communication between Defra and NAES

There is currently an agreed service level agreement (SLA) between National Accounts and Economic Statistics (NAES) Group and Defra for the supply of monthly, quarterly and annual data for the 11 components ONS requires.

The SLA has been in place since 2015, is a 3-year rolling contract until 2018 and provides a formal agreement between the two departments as to where responsibilities and requirements lie.

Defra is required to:

- make NAES aware of any changes in their methods
- provide advance warning if unable to provide data
- provide guidance on fluctuations in data and advice on changes with weights, that is, Defra should provide detailed briefings
- Defra are responsible for ensuring that series are consistent over time; where this is impossible to achieve, Defra should then agree a course of action with NAES to achieve a satisfactory series; p-rocedure would be for Defra to discuss a way forward with NAES branch head.

Defra have been asked to inform NAES of any factors that may be helpful in better informing estimates whenever these become available but especially during November of each year when forecasts for the coming year are reviewed in the spreadsheet system (section 3.4).

The SLA states that any concerns NAES has with data will be queried with the main point of contact in Defra named on the SLA. The records of queries will be kept and maintained within NAES.

However, it should be noted that there have been changes in roles for both Defra and NAES and the contacts details on the SLA need to be updated.

The current method of communication between NAES and Defra is via email when new data is being sent and there are currently no regular formal meetings in place. During recent discussions between NAES and Defra about the SLA it was discovered that the change in methodology for potatoes had not been communicated to NAES, which is a requirement within the SLA. Furthermore the contact information and the specific wording of the SLA need to be updated. There is now ongoing discussion to update the SLA, which will prevent such occurrences in the future. Formal annual meetings will be scheduled to discuss the SLA. This shows that whilst there was a failure in communicating the changes to potatoes, Defra are committed to improving their communication with NAES.

Strengths

- survey forms sent accompanied by further information and guidance with helpline available
- Defra chase up non-responders or query suspect data
- there is a service level agreement in place between NAES and Defra formally detailing responsibilities
- Defra regularly communicate with their data suppliers

Weakness

- changes to the methods for potatoes were not communicated to NAES
- contact details on the SLA are out of date
- no formal meetings between Defra and NAES

Next steps

- communicate on a more regular basis including the arrangements of formal meetings
- updating aspects of the current SLA, for example, contact details

3.3 Quality assurance principles, standards and checks by data supplier – matrix score, A2

This relates to the validation checks and procedures undertaken by the data supplier, any process of audit of the operational system and any steps taken to determine the accuracy of the administrative data.

A process map showing the quality assurance checks conducted by Defra is shown visually in Figure 4 of the annex.

3.3.1 Quality assurance Checks for all of the components

The data for wheat, barley, oilseed rape, cattle, sheep, pigs, poultry, egg, milk and horticulture are subject to a variety of rigorous validation checks, by Department for Environment, Food and Rural Affairs (Defra). This includes:

- automatic checks
- historical checks
- manual feedback
- user feedback, which provides an independent aspect to identify inconsistencies in the data such as outliers and extreme values

Automatic checks involve using software to identify outliers and/or anomalies in the data. Historical checks entail comparing respondents' survey results from previous years in order to assess whether there are any outliers. Manual checks involve contacting the survey respondent where the new data seems unusual in comparison with previous years.

Using the case study example for cattle, sheep and pigs – prior to publishing the statistics they are subject to scrutiny by two main users:

Eurostat/ The EU Commission where these slaughtering statistics are required monthly under Regulation number EC. 1165/2008. Defra send the data to Eurostat, which includes briefings and commentary on the data. This gives Eurostat an opportunity to scrutinise the data before Defra publish on their website.

The livestock industry, this includes divisions of the Agriculture and Horticulture Development Board (AHDB), AHDB Pork 16 and AHDB Beef and Lamb . Users of slaughtering statistics rely heavily on the statistics in order to assess the current state of the industry and predict the available supplies of meat production for the coming year. This can then affect meat prices and trade decisions on levels of imports and exports to maintain supply. Defra work closely with AHDB and prior to publication will engage with them to discuss if the data are behaving strangely.

After the statistics have been published there is further opportunity for Defra's wider users to give feedback and scrutinise the results. Defra have stated that the data has not been strongly challenged and there has been no feedback querying the results.

As the statistics are rarely challenged by Defra's users and the general public, Defra are confident that the quality assurance checks meet users' needs and are fit for purpose.

For potatoes there is a general quality assurance based on reports in farming press and other details published on the web such as the AHDB website.

Strengths

- Defra have a variety of rigorous quality assurance checks in place
- survey results can be compared with other data sources, for example, from the Food Standards Agency
- data are discussed with main industry players and users prior to publication

Weaknesses

- Defra have to adhere to a tight deadline and therefore publish within 3 weeks of receiving the data

3.4 Producers' quality assurance investigation and documentation – matrix score, A1

This relates to the quality assurance conducted by the statistical producer, including corroboration against any other data sources.

3.4.1 Quality assurance by National Accounts Economic Statistics (NAES) Group

A process map showing the quality assurance checks conducted by output producers is shown visually in Figure 5 of the annex.

As mentioned in Table 3, Department for Environment, Food and Rural Affairs (Defra) supply National Accounts and Economic Statistics (NAES) Group with 11 components of data. When the data is received it is quality assured (QA) by a member of the NAES team who is the main point of contact for agriculture. The monthly, quarterly and annual data are delivered via a spreadsheet and includes new data as well as any revisions to previous data. Where the data contains anything unusual or unexpected Defra provide briefings to explain this.

NAES then conduct the following quality assurance checks:

- compare new data with the previous period data and assess growth trends
- compare new data with the same period in the previous year (that is, January 2017 to January 2016) to assess historical, seasonal trends and growth trends
- carry out sense checks on submitted data, for example, check totals add up

On those occasions when any of the monthly, quarterly and annual data are not available in time for publication NAES make forecasts. For example, NAES will publish in April but the quarterly and annual data will not be available so forecasts have to be made. In this case data imputation is built into the system that forecasts a figure based on historical data and trends from previous years.

After the data has been quality assured, the NAES team then deflate and seasonally adjust the data. After this process there are no further quality assurance checks. The data passes through an internal ONS system for the final output where the agriculture data (division 01) is aggregated with divisions 02 forestry and 03 fishing to produce part of the GDP(O) estimate.

NAES also benchmark to Defra annual supply and use data when this is made available to ensure that there is consistency between Defra and NAES. This conceptually superior data inevitably leads to revisions to final outputs but these do not significantly impact NAES estimates of GDP. Defra are keen to work with NAES to gather feedback to improve future data sources.

There are full desk instructions in place to carry out these functions, however, this is potentially rather complex for a new starter. The NAES team are currently in the process of reviewing and updating these desk instructions. It should be noted that despite there being complex desk instructions there are a number of experienced members in the team who can provide guidance and expertise. The expertise within the team and current desk instructions would enable another person in the team to pick up the quality assurance processes and sign them off.

Strengths

- good quality assurance checks in place
- built on already rigorous checks conducted by Defra
- no major feedback or concerns raised from NAES users about agriculture

Next steps

- review and update desk instructions

Notes for : Areas of quality assurance

1. Data for Scotland are not available separately due to confidentially reasons.
2. Northern Ireland sends the data to DEFRA and there is limited detail published for milk.
3. Department for Environment, Food and Rural Affairs (Defra) define major based on throughput.
4. The Food Standards Agency (FSA) do not currently publish their data. They send Defra a copy of the data each month.
5. This is the livestock slaughtered for meat consumption.
6. For more information about what guidance Defra provide to their data supplier when completing the surveys see section 3.2.
7. This acts states that Ministers must publish an annual report relating to price support for agricultural produce as they consider relevant. The government uses this information when considering policy issues such as proposals by the European Commission in regards to the Common Agricultural Policy (CAP) and the provision of agricultural support.
8. That is, they were not empty properties or business addresses.
9. Defra have confirmed that they are in the process of producing a methodology paper for the production of potatoes in the near future but do not currently have a timeframe.
10. Due to confidentiality reasons Defra are unable to provide details of these key industry contacts.
11. These factors were estimated based on commissioned survey and research in 2003 to 2004 and then updated in 2005 to 2006.
12. Farm-gate prices (in £ per tonne) or value based on farm-gate price. This is the price the farmer is paid for their produce with no extra delivery or packaging costs.

4 . Summary

A process map showing the overall flow of agriculture statistics from suppliers to National Accounts Economic Statistics (NAES) Group is shown visually in Figure 6 of the annex. In investigating the agricultural data source NAES consider the main strengths of the data to be:

- 90% to 100% of industry coverage
- many of the surveys are collected by law
- Defra use a variety of rigorous checks to quality assure the data such as automatic and historical checks, manual and user feedback
- NAES quality assure the data by comparing growth rates from the previous period as well as comparing the same time period in previous years
- there is an up to date service level agreement (SLA) in place

The current limitations of this data source are:

- the change in methodology for potatoes had not been communicated to NAES
- the SLA is not up-to-date with out of date contacts and incorrect wording
- shortcomings in the estimations of horticulture

In constantly seeking to improve our data sources we will be taking next steps to address these limitations and these will be communicated to users in the future quality report updates for this topic.

However, despite these slight limitations based on the medium risk of quality concerns and contribution that the agriculture statistics feed into gross domestic product (GDP) (0.67%) NAES consider this data source to fulfil the requirements of an A2 assurance rating.