



United States Department of Agriculture
Office of Inspector General





United States Department of Agriculture
Office of Inspector General
Washington, D.C. 20250



DATE:

AUDIT
NUMBER: 50601-0015-KC

TO: Dr. Cynthia Clark
Administrator
National Agricultural Statistics Service

ATTN: Michelle Garner
Acting Director
Financial Management Division
Agricultural Research Service

FROM: Gil H. Harden
Assistant Inspector General
for Audit

SUBJECT: National Agricultural Statistics Service Establishment of Average Yields

This report presents the results of the subject audit. Your written response, dated April 8, 2011, to the official draft report is included, in its entirety, at the end of this report. Excerpts from your response and the Office of Inspector General's position are incorporated into the relevant sections of the report.

In the written response to the official draft report, the agencies concurred with the findings and recommendations in the report. Based on your response, we have reached management decision on all recommendations.

In accordance with Departmental Regulation 1720-1, final action needs to be taken within 1 year of each management decision to prevent being listed in the Department's annual Performance and Accountability Report. Please follow your internal agency procedures in forwarding final action correspondence to the Office of the Chief Financial Officer.

We appreciate the courtesies and cooperation extended to us by members of your staff during our audit fieldwork and subsequent discussions.

Table of Contents

Executive Summary	1
Background & Objectives	5
Section 1: County Yield Estimates	7
Finding 1: NASS Underutilizes Data Available from RMA To Estimate County Yields.....	7
Recommendation 1: To NASS.....	10
Recommendation 2: To NASS and RMA.....	10
Recommendation 3: To NASS and RMA.....	10
Finding 2: FSA Needs Controls To Ensure the Accuracy of Data in the Crop Acreage Data Mart.....	12
Recommendation 4: To FSA.....	13
Recommendation 5: To FSA and NASS.....	14
Scope and Methodology.....	15
Abbreviations	16
Agencies' Response	17

National Agricultural Statistics Service Establishment of Average Yields

Executive Summary

Among other benefits, the U.S. Department of Agriculture's (USDA) National Agricultural Statistics Service (NASS) provides national, State, and county crop estimates, which USDA agencies such as the Risk Management Agency (RMA) and the Farm Service Agency (FSA) use to determine program benefits.¹ For example, RMA's area-based (county average) crop insurance indemnities—totaling nearly \$1.1 billion from 2007-2009—are based partly on NASS' county-level crop yield estimates. Since NASS' estimates affect program benefits, we undertook this audit to determine: (1) whether the processes that NASS uses to estimate county production yields generate accurate and reliable information, and (2) whether the data collected to make these estimates offer a reasonable basis for administering programs.

Overall, we concluded that NASS' current methodology for estimating county yields does provide reasonably accurate and reliable information. During the course of our audit, we identified opportunities for NASS to improve its data collection (and subsequent county yield estimates) by using RMA's information to supplement and corroborate its own, and by entering into formal data-sharing agreements with RMA and FSA.

NASS' process for estimating crop acreage and production is complex, but essentially the agency collects the necessary data from producers through surveys. NASS' statisticians then analyze the responses and establish estimates. Since NASS' methodology is geared to provide national and State estimates, a single county's information does not necessarily affect these broader estimates.²

If enough producers do not respond to its surveys, NASS does not publish average yields for these counties. Since 2008, the agency began requiring (with limited exceptions) at least 30 responses to estimate a county's yield, which resulted in fewer county estimates for fewer crops.³ After establishing the more stringent standard, for example, the number of NASS' county-level corn production estimates dropped from 1,896 to 1,551 over 3 years (2007-2009)—an 18-percent decrease.⁴ For Federal programs such as RMA's area-based Group Risk Plan (GRP) and Group Risk Income Protection (GRIP) insurance programs, which are based on a county's average production to set payments, the decrease in county estimates has had a negative impact because RMA is required to withdraw future offers of insurance coverage to

¹ Producers also use NASS' estimates of crop yields to make planting and marketing decisions. In addition, agricultural organizations, trade groups, financial institutions, and other entities rely on the agency's work for purposes ranging from pricing commodities to storing them.

² According to NASS, its national and State estimates are more statistically sound than county-level estimates. NASS acknowledged that county-level estimates are less reliable due to a more limited number of responses.

³ In cases where there are fewer than 30 responses from an individual county, NASS is allowed to publish an estimate for "combined counties."

⁴ To determine this, we downloaded and analyzed NASS estimates of total yield (per crop, per county) for corn using a historical 10-year average (e.g., crop years 1998 through 2007, and separately for crop years 2008 and 2009).

producers in these affected counties, thereby reducing overall opportunities to participate in these programs.

Without NASS county-level estimates, RMA has not been able to offer these insurance plans in affected counties. For example, in a December 2009 program announcement, RMA stated that it would terminate 469 corn county crop programs in 2010 due in part to NASS' more stringent standards.⁵ For many such counties, though, RMA has a wealth of data to supplement NASS' own, which would help generate estimates so producers can participate. For example, RMA-underwritten insurance programs have a high participation level—84 percent of NASS-estimated corn acreage and 82 percent of NASS-estimated soybean acreage—because the major corn and soybean producing States are RMA-reinsured.

NASS can use the data RMA collects from these participants to help generate estimates in counties that did not have sufficient survey response rates. In turn, RMA and other agencies can use these estimates to offer insurance or other program benefits.

NASS can also use RMA's data to corroborate county estimates. Based on our analysis, RMA's crop yield data are often fairly close to NASS' estimates. Therefore, we believe that RMA's data can be used to help NASS' statisticians cross-check county-level estimates.⁶ For example, about 80 percent of NASS' county estimates and RMA's collected data were generally within 10 percent of each other for corn and soybean yields between 2006 and 2008.

Accordingly, significant deviations between NASS' estimates and RMA's data may indicate that NASS needs to cross-check one with the other. For example, in 2008, NASS' corn estimates in one Colorado county were 88,000 acres harvested and 7.2 million bushels produced. After NASS considered RMA data for this county, NASS lowered its estimates for acres harvested and production, decreasing them to 48,600 acres and 5.23 million bushels—45-percent and 27-percent reductions.

Currently, RMA provides NASS information on a limited basis according to a Memorandum of Understanding between the agencies.⁷ Typically, NASS receives RMA's acreage data in December. NASS officials have indicated, though, that they are reluctant to use other agencies' information due to differences in data definitions⁸ that are involved with moving data from one database to another. On the other side, RMA has stated that giving NASS unrestricted access to its data will create a burden on its system. However, we believe that any additional demand on

⁵ RMA announced over 1,000 county crop programs were removed in 2010 in the December 1, 2009, program announcement.

⁶ NASS does use some indicators to evaluate its estimates, such as comparing the current year to the previous year's estimates, satellite indicators, and harvested to planted ratios. However, these indicators do not fully account for variable growing or extreme weather conditions that may have existed during the survey period.

⁷ A Memorandum of Understanding effective January 1, 2008, between RMA and NASS outlines the agencies' general data sharing agreement. The Memorandum of Understanding does not provide for allowing NASS full ongoing access to RMA's automated database systems which is needed in order for NASS to take full advantage of acreage, yield, and loss data contained in RMA's automated system.

⁸ For example, NASS does not count corn silage as production for grain even though production information reported to RMA may not distinguish the difference between corn harvested for grain and corn harvested for silage.

RMA's system could be limited to one NASS headquarters official obtaining system access. We acknowledge these issues, but note that in the absence of a formal agreement where RMA specifically grants full ongoing access to NASS, they have not yet worked out which issues can be resolved and which cannot. Accordingly, we recommend that, to the extent practicable, the two agencies develop a formal agreement or augment the current Memorandum of Understanding that maximizes NASS' ongoing access to RMA's data.

During the course of our audit, we learned that duplicate records in FSA's crop acreage database⁹ led to NASS overstating crop estimates in October 2008. We traced this issue to a lack of oversight and data validation by FSA. On February 26, 2009, the Office of Inspector General (OIG) issued a memorandum to the Deputy Under Secretaries for Research, Education, and Economics and Farm and Foreign Agricultural Services on problems with the FSA data that NASS used for its October 2008 Crop Production report. Specifically, NASS utilized the crop acreage information obtained from FSA's "Crop Acreage Data Mart" to establish planted acreage estimates for certain crops published in the October 2008 Crop Production report. In our memorandum, we recommended that NASS document the process for using data from other agencies in its Crop Production reports. We also recommended that FSA improve its internal controls in order to eliminate the identified deficiencies in its Crop Acreage Data Mart system. Both agencies responded that they had completed the recommended actions.

Recommendation Summary

To NASS and RMA:

- Establish a formal agreement or augment the current Memorandum of Understanding giving NASS ongoing access to RMA data to the extent practicable given the systems and the agency's reporting requirements.

To NASS and FSA:

- Establish a formal agreement between the agencies that institutes the controls necessary to ensure the integrity of data needed with respect to their intended use.

To NASS and RMA:

- Implement the agreement by undertaking those tasks necessary by each agency to enable NASS full access to RMA data.

To FSA:

- Complete corrective actions to improve FSA's internal controls in order to eliminate the identified deficiencies in the agency's Crop Acreage Data Mart system.

⁹ NASS relies on FSA crop acreage data for crop production reports it publishes in September and October to ensure that planted acres reported to FSA are covered in NASS' estimates.

Agencies' Response

NASS provided a combined response to the official draft report for the three agencies. In their April 8, 2011, response to the official draft report, NASS, FSA and RMA agreed with the findings and recommendations and provided information on corrective actions planned and underway.

OIG Position

Based on the information contained in the combined response, we accept management decision for the five recommendations in this report.

Background & Objectives

Background

NASS is USDA’s primary statistical agency and its mission is to provide timely, accurate, and useful statistics in service to U.S. agriculture. The agency issues about 425 statistical reports each year, covering virtually every facet of agriculture at the county, State, and national levels. These reports include data on 120 crops.

All NASS-developed county estimates derive from data collection programs through cooperative agreements, mostly with State departments of agriculture. States are willing to cooperate with NASS in exchange for data describing their agricultural economies at the county or district (i.e., multi-county) level.

NASS’ annual process to develop estimates for average crop acreage and production is complex. For each of certain predetermined crops—selected for each State depending on the region and which crops predominate—NASS uses producer-reported data to estimate the Statewide totals for the number of acres planted, the number of acres harvested, and total production from the harvested acres. NASS obtains these data from survey questionnaire responses, computer-assisted telephone interviews, in-person interviews conducted by NASS field enumerators,¹⁰ and online reports from producers.

NASS procedures require use of “probability-based surveys,” which are based on random samples chosen using statistically sound methodology, and are thus associated with a given confidence level. NASS’ September and December Agricultural Surveys, which are probability-based surveys, serve as the primary basis for determining and publishing the national and State acreage and production estimates each year in January, for the prior year’s crop.¹¹ These published State estimates establish a benchmark; that is to say, when NASS determines county-level estimates for each State and crop, the total county-level acreage and production estimates for a given State must equal the published State-level numbers. If NASS adjusts estimates in one county, it will include an offsetting adjustment in another county or counties to maintain the overall State benchmark.

However, NASS also utilizes surveys that are not probability-based. They fall into the category of “non-probability-based surveys.”¹² These two late-season surveys, the County Estimates Survey (CES) and the County Agricultural Production Survey (CAPS), which are conducted by NASS field offices at the county level any time from September through November each year, provide the majority of survey data needed to construct the county-level estimates that NASS disseminates the following March.¹³ These two surveys essentially cover the major crops grown

¹⁰ Enumerators can be information gatherers or census takers, etc.

¹¹ The September Agricultural Survey examines summer-harvested “small grains” (wheat, oats, and barley), and the December Agricultural Survey examines “row crops” (corn, soybeans, sorghum, dry beans, sunflower, alfalfa, forage crops, and hay).

¹² A non-probability survey is defined as “a survey that does not conform to the definition of probability survey.” The sample is not randomly selected, but instead is selected based on a specific methodology.

¹³ The survey time-frames vary because of variable harvest times and State field office workload.

and livestock raised by producers in a given area for a given year, while the Quarterly Agricultural Surveys are limited to specific crops.

After receiving survey responses, enumerators from the National Association of State Departments of Agriculture or statisticians from NASS subject them to edit and quality controls.¹⁴ When producers provide ambiguous responses, NASS may make follow-up telephone calls. From this point, accumulated data flow through the Interactive Data Analysis system, which allows NASS officials to view all the data in graphic format.¹⁵ This county-level data is composed of both CES/CAPS data and data from the September and December Agricultural Surveys. The Database Integrated County Estimates (DICE) system¹⁶ is the primary data system that NASS uses to store the producers' survey results.

NASS statisticians then add other sources of county-level information, such as estimates from prior years, periodic census-level data, remote-sensing data,¹⁷ and FSA/RMA operational data, to the dataset. They summarize the combined dataset and develop composite indications/ratios to help them determine the most likely production and acreage numbers. They scale the data back to district and State numbers, and perform automated computer checks to ensure that the data are consistent and balanced across the State, and from county level to State level.¹⁸

The preliminary estimates are then reviewed—first at the State field office level, and subsequently by NASS headquarters officials—before being finalized and disseminated.

Objectives

Our overall objectives were to assess: (1) whether the processes that NASS uses to estimate county average production yields provide accurate and reliable information, and (2) whether the data collected to make these estimates provide a reasonable basis for administering programs.

¹⁴ The edit and quality control work is done through the Blaise system, which is a Windows-based computer-assisted telephone interviewing system and survey processing tool. NASS uses Blaise software for the production of official statistics; the software provides NASDA enumerators and NASS field office personnel (specifically, NASS statisticians) the ability to interactively review and edit incoming survey results.

¹⁵ The Interactive Data Analysis system is a tool designed to aid NASS field office statisticians and management to graphically analyze the survey data set for distributions of acreage or yields by county or by district. Extreme outliers and influential reports can then be examined for feasibility. Also, NASS statisticians and management can compare one year's data to another year's, and can examine the ratio of planted acreage to harvested acreage.

¹⁶ DICE is the primary repository for data acquired in NASS surveys. NASS has adopted it in stages. DICE has been in use since 2007 for only a few major crops, was expanded to include most crops in 2008 and 2009, and will likely include all crops for the 2010 crop year.

¹⁷ Remotely sensed county estimates are created by supervised classification of satellite images and use of small area estimation techniques to produce statistically valid estimates of acreage.

¹⁸ Scaling is a weighting adjustment of preliminary county-level survey data in order to account for any incompleteness due to low survey response rates, historical indicators, or administrative data. Scaling works through mathematical adjustments of county-level estimates and indicators (survey-based direct expansions, estimates of ratios (like the ratio of harvested acres to planted acres), and estimates from previous years), to ensure that all counties in a given district add up to the district total and that all districts in a given State add up to the State estimated total, proportionately.

Section 1: County Yield Estimates

Finding 1: NASS Underutilizes Data Available from RMA To Estimate County Yields

NASS is not fully using RMA operational data to supplement and corroborate the volunteered, producer-reported information that NASS employs to estimate county crop yields.¹⁹ In addition to providing the opportunity to cross-check NASS' own information, RMA's data offer NASS information about producers who do not respond to NASS' voluntary surveys. NASS has hesitated to use this other source of information more fully because it considered RMA's reporting timeframes as out of step with its own county estimate publishing dates. In addition, NASS is reluctant to use RMA's data because of differences in data definitions that can arise when moving information from one system to another. As a result, NASS has not taken full advantage of the opportunity both to enhance its county-level estimation process and to strengthen its interagency coordination through improved data-sharing.

NASS' mission is to provide timely, accurate, and useful statistics in service to U.S. agriculture. The agency is to provide objective and unbiased statistics on a preannounced schedule that are fair and impartial to all market participants. By implication, NASS' county-level statistics need to be as accurate as possible for the agency to achieve this goal.

The impact of slightly errant NASS' county level estimates is significant because of the magnitude of program benefits at risk. For example, we examined counties where NASS' estimated yields differed from RMA-certified yields by greater than plus or minus 10 percent. Through that analysis, we determined that for corn, about 20 percent of NASS' estimated yields differed from RMA-certified yields by greater than plus or minus 10 percent for 346 counties, from 2006 through 2008. For soybeans, about 15 percent of NASS' estimated yields differed from RMA yields by the same 10 percent, for 242 counties, from 2006 through 2008.²⁰

For those corn and soybeans counties outside the plus or minus 10 percent difference, GRP/GRIP program liabilities totaled \$366.5 million, premiums were \$30 million, and indemnities totaled \$23.3 million. Granted, while there is no reason to support one USDA agency's county-level crop estimates or aggregated farm-level data over the other, the difference indicates a need for further analysis and corroboration, particularly when large amounts of program payments are involved.

In addition to helping cross-check estimates, RMA's data provides NASS with important information when producers do not respond to NASS' surveys.²¹ The response rates to NASS'

¹⁹ The term "operational data" includes data on the following: planted acreage, failed acreage, unharvested acreage, prevented planting acreage, harvested acreage, and harvested production. In general, producers have certified the accuracy and completeness of such data. The data is also subject to spot check and review by the administering agency—in this case, RMA.

²⁰ For corn we identified 159, 97, and 90 counties, and for soybeans we identified 101, 81, and 60 counties from 2006 through 2008, respectively.

²¹ Reporting acreage and production information to NASS is voluntary.

surveys typically range between 60 and 75 percent.²² RMA's data provide an opportunity to supplement NASS' county level data, especially where county response rates are low, as can happen, for example when natural disasters and/or extreme weather conditions coincide with attempts to solicit survey responses.²³

Furthermore, RMA's data can help NASS develop estimates for counties where the response requirements are not met. Since 2008, NASS has required at least 30 responses from a county to generate a crop yield estimate for it.²⁴ For example, after establishing the more stringent publication standard, the number of NASS' county-level corn production estimates decreased from 1,896 counties to 1,551 counties from 2007 to 2009—an 18-percent decrease.²⁵ For soybeans over the same period, there was a similar decrease in NASS' county estimates, which fell from 1,509 counties to 1,382 counties from 2007 to 2009—an 8-percent decrease.

Without the availability of NASS county-level estimates, RMA has been unable to offer insurance plans such as GRP and GRIP²⁶ for some counties, stating in December 2009 that it would terminate over 1,000 county crop programs in 2010 due in part to NASS' more stringent 2009 standards.

RMA's data offer a supplemental source of information that could be sufficient to allow NASS to generate an accurate estimate in some cases.²⁷ Because RMA-underwritten insurance programs have a high participation level, RMA has a tremendous amount of data on acreage and production. For example, 84 percent of NASS-estimated corn acreage and 82 percent of NASS-estimated soybean acreage in the major corn and soybean producing States are RMA-reinsured.

In addition, RMA's crop yield data are often fairly close to NASS' own estimates. Our analysis of over 500 counties in 11 States showed that for the 2006 through 2008 corn crop, about 80 percent of the county yields we derived from RMA's production and yield data were within 10 percent of NASS' estimates for the corresponding counties. For 2006 through 2008 soybean crop, about 85 percent of the RMA-derived county yields were within 10 percent of NASS' estimates for the corresponding counties.

²² While NASS can complete the data universe (i.e., impute missing data from non-responders) with its probability-based State and national surveys, it cannot do so for county surveys because they are not probability based—i.e., not based on a random sample chosen using statistically sound methodology.

²³ NASS can use its resources and means to cross-check this information in aggregate.

²⁴ In cases where there are fewer than 30 responses from an individual county, NASS is allowed to publish an estimate for “combined counties.”

²⁵ To determine this, we downloaded and analyzed NASS estimates of total yield (per crop, per county) for corn for crop years 2007, 2008, and 2009.

²⁶ Title 7 *Code of Federal Regulations* 407.9, dated January 1, 2009, requires RMA to use county yields set by NASS to determine indemnities for policies covered under the GRP/GRIP plans of insurance.

²⁷ Producers with crop insurance coverage certify their production annually to their Approved Insurance Providers—the private-sector entities reinsured by RMA's Federal Crop Insurance Corporation—who then report these data to RMA.

In these counties, RMA data “covered” at least 40 percent of NASS’ estimates of the planted acreage.²⁸ The depth, coverage, and accuracy of RMA’s data offers NASS an opportunity to both supplement and corroborate its own county-level estimates when needed.

For example, NASS’ corn estimates in one Colorado county were 88,000 harvested acres, and 7.2 million bushels produced. After NASS considered RMA data for this county, NASS lowered its estimates for acres harvested and production, decreasing them to 48,600 acres harvested and 5.23 million bushels of corn produced—reductions of 45 percent and 27 percent, respectively.

As it collects data, NASS does make some use of RMA’s crop acreage information.²⁹ NASS also receives crop acreage data from RMA each December—information on prevented planting acres, unharvested acres, and harvested acres with insurance claims. While NASS is provided RMA’s acreage data in December for the most recent crop year, the corresponding yield information has just started to be collected by RMA and NASS does not have full access to RMA yield data. NASS officials stated that December is the optimal timeframe for using these data, since any later time would make it difficult to integrate the information into the county estimation process. We note, though, that by mid-February, about 35 percent of RMA’s yield data for that year is available and could be used, at least partially, to supplement and corroborate where appropriate.

NASS and RMA currently operate under a Memorandum of Understanding that covers the exchange of basic information between the agencies. NASS officials have also indicated their reluctance to use other agencies’ information due to differences in data definitions³⁰ and, consequently, issues inherent in moving data from one system to another. In addition, RMA has concerns that giving NASS unrestricted access to its data will create a burden on its system. However, as of May 2010, the two agencies have begun to work together and investigate the timeliness of RMA yield data received during the late winter/spring months and to potentially work out a practical method for RMA to provide these data to NASS each year. We acknowledge these limitations, but note that in the absence of a formal agreement where RMA specifically grants full ongoing access to NASS, or augmenting the current Memorandum of Understanding, they have not yet worked out which issues can be addressed and which cannot.

Accordingly, we recommend that, to the extent practicable, the two agencies develop a formal agreement or augment the current Memorandum of Understanding that maximizes NASS’ full, ongoing access to RMA’s data. We understand that the ultimate scope of that agreement will be determined partly by such factors as the nature of the systems involved, the intended use, and required timing for the data.

²⁸ Since NASS makes estimates for the total amounts of acreage, production, etc., the real-life data from RMA will necessarily reflect—or “cover”—only a subset of NASS’ data universe. For example, in 2009, RMA insured corn acreage in Illinois “covers” about 81 percent of NASS planted acres.

²⁹ FSA made crop acreage information available to NASS through its Crop Acreage Data Mart, which has been shut down until data quality issues are addressed (see Finding 2).

³⁰ For example, NASS does not count corn silage as production for grain even though production information reported to RMA may not distinguish the difference between corn harvested for grain and corn harvested for silage.

Recommendation 1: To NASS

Develop a strategy to use RMA acreage, production, and yield data to cross-check and supplement NASS' process for generating county-level yield estimates.

Agency Response

NASS officials concurred with this recommendation and will explore the use of RMA data and develop a strategy to utilize these data in the most appropriate fashion for NASS county-level yield estimates.

This exploration and plan on utilizing the data will be conducted by NASS during the 2011 production season. A targeted completion date for the review and resulting plan is April 1, 2012.

OIG Position

We accept management decision for this recommendation.

Recommendation 2: To NASS and RMA

Establish a formal agreement granting NASS full, ongoing access to RMA data to the extent practicable given the systems and the agency's reporting requirements.

Agencies' Response

NASS and RMA officials concurred with this recommendation and have established a formal agreement giving NASS ongoing access to RMA data. This access is formalized in an Interconnection Security Agreement between RMA and NASS effective on December 14, 2010.

OIG Position

We accept management decision for this recommendation.

Recommendation 3: To NASS and RMA

Implement the agreement by undertaking those tasks necessary by each agency to enable NASS full access to RMA data.

Agencies' Response

NASS and RMA officials concurred with this recommendation and, through an Interconnection Security Agreement between the agencies effective December 14, 2010, have undertaken the tasks necessary for each agency to perform in order to enable NASS full access to RMA data.

OIG Position

We accept management decision for this recommendation.

Finding 2: FSA Needs Controls To Ensure the Accuracy of Data in the Crop Acreage Data Mart

On October 24, 2008, the then Under Secretary for Research, Education, and Economics notified OIG of a problem with NASS' October 2008 Crop Production Report, published on October 10, 2008, and indicated that a revised report would be issued.³¹ NASS subsequently issued a corrected October 2008 production report on October 28, 2008. Based on the information provided us, we decided to initiate a review of the conditions that led NASS to revise the report.

Accordingly, we reviewed the processes that NASS uses to develop crop acreage estimates for its crop production reports. When NASS established planted acreage estimates for certain crops published in the September³² and October³³ crop production reports, the agency used the crop acreage information obtained from FSA's Crop Acreage Data Mart.³⁴ It turned out that the data mart contained duplicate acreage records, causing some of the planted acreage and crop production data to be overstated.

The data errors in the data mart and the resulting errors in the October 2008 production report came to light after mid-October 2008, when two former USDA employees contacted current FSA Economic and Policy Analysis staff, questioning the corn and soybean acreages in the report. This contact spurred FSA's staff to review the data mart's information. Specifically, by cross-checking the crop acreage totals, they found that the data mart contained duplicate records.

FSA was unable to determine when duplicate records began appearing and how many there were because the data mart is continuously updated but does not always overwrite farmer data. During the updating process, some farmer data were loaded multiple times, causing duplicate records. Since there was no ownership of this data system by any FSA unit, these duplicate records were not identified as such and rejected by the system. As a result, multiple records for the same producers appeared in the data mart, causing acreage information to be overstated. FSA has suspended the use of the data mart while it determines how to correct the identified problems.

We concluded that FSA has not maintained adequate system controls to ensure the accuracy of the information in the data mart. FSA developed the system, but did not assign any division with specific responsibility for maintaining it. FSA also did not incorporate business logic into the

³¹ Since 2003, NASS has been accessing FSA's crop acreage information (certified by the producers) and factoring this information into its Crop Production Reports. The Crop Production Reports are important; producers, agricultural organizations, businesses, trade groups, and financial institutions use them when making planting, pricing, and policy decisions, and thus they impact not only USDA's agricultural programs but also business decisions and markets.

³² Each September, NASS updates the number of acres planted for cotton, rice, and peanuts, using FSA data on the minimum number of acres planted because these data are reported and certified by producers who participate in FSA programs.

³³ Each October, NASS updates the number of acres planted for dry edible beans, canola, corn, sorghum, soybeans, and sunflowers, using FSA data on the minimum number of acres planted because these data are reported and certified by producers who participate in FSA programs.

³⁴ FSA's Crop Acreage Data Mart provides the FSA national and State offices and NASS with the capability to obtain data from the Forms FSA-578, "Report of Acreage."

agency's programs (i.e., including within the data mart different codes for specific cropping practices). In addition, FSA has not implemented sufficient application controls, such as those necessary to prevent duplicate records.

In our February 2009 memorandum to the Deputy Under Secretary for Research, Education, and Economics and the Deputy Under Secretary for Farm and Foreign Agricultural Services, we recommended that FSA take corrective actions to improve its system controls in order to eliminate the identified deficiencies in its Crop Acreage Data Mart. Specifically, we recommended that FSA:

1. develop application controls to identify and prevent duplicate records,
2. incorporate accurate business logic into the data mart to keep the system consistent with program changes,
3. assign one of FSA's internal units the specific responsibility for maintaining the data mart, and
4. validate the information contained in the data mart to ensure its accuracy on a regular basis.

In June 2009, FSA provided a response to OIG, stating that the agency has eliminated the deficiencies identified in its data mart through multiple additional controls. However, at that time, the system remained inaccessible for NASS' use. In addition, NASS provided a response in January 2010, stating that it will promote data transparency through internal documentation of the use of FSA administrative data, and through public documentation of the use of administrative data used in the estimation process.

Moving forward, a formal agreement between the two agencies will improve NASS' ability to identify and mitigate issues that can develop from interagency data-sharing and FSA's ability to suit its controls over data integrity to meet NASS' needs.

Recommendation 4: To FSA

Provide evidence that the corrective actions to address the data mart's deficiencies have been completed, including:

- a. application controls to identify and prevent duplicate records;
- b. accurate business logic incorporated to keep the system consistent with program changes;
- c. division assigned responsibility for maintaining the data mart; and
- d. data mart's information validated on a regular basis.

Agency Response

FSA officials concurred with this recommendation and have completed development of the Crop Acreage Data Warehouse (CADW) changes required to implement the corrective measures recommended by OIG. CADW is in the final certification testing phase prior to re-deployment. A team of FSA State and national office and NASS employees have been identified to test the

data mart. Corrective measures will be verified by the certification testing effort. Once the verification is complete the data mart will be re-deployed to production.

In order to prevent duplicates from occurring, detail acreage reporting data with primary keys were added to CADW. This approach also allows detailed verification of data synchronized with the mainframe acreage reporting data.

The AS/400, mainframe, and data mart files are monitored and synchronized daily. Control totals are contained within the daily transmission files and are balanced to the mainframe data. Exception reports are created and monitored daily. These exceptions are researched. Upon completing the research, if required, a full file upload is triggered in order to synchronize the acreage data.

The Production, Emergency, and Compliance Division is now the business owner for CADW. FSA's Database Management Office along with the Production Adjustment and Risk Management Office provide IT support for CADW.

Subject matter experts from FSA's Database Management Office and the Production Adjustment and Risk Management Office are assigned to maintain the integrity of CADW by reviewing daily changes and assessing the impact. Starting with Crop Year 2011, acreage reports are being captured in a centralized web based system. This simplifies validation as it eliminates synchronization of the data between the AS/400 and the mainframe databases that feed CADW. FSA subsequently advised us that CADW will be redeployed in September 2011.

OIG Position

We accept management decision for this recommendation.

Recommendation 5: To FSA and NASS

Establish a formal agreement between the agencies to institute the controls necessary to ensure the integrity of data provided with respect to their intended use.

Agencies' Response

FSA and NASS officials concurred with this recommendation and will revise the existing Memorandum of Understanding to address the controls necessary to ensure the integrity of the data with respect to its intended use. The Memorandum of Understanding will be revised by September 30, 2011.

OIG Position

We accept management decision for this recommendation.

Scope and Methodology

We performed our audit from May 2008 through December 2010 at NASS, RMA, and FSA Headquarters in Washington, D.C.; RMA's Product Management Office in Kansas City, Missouri; two NASS State field offices (Iowa and Colorado); and one FSA county office (in Colorado). Our review covered NASS-established county average yields used by USDA agencies to determine program benefits for crop years 2006 through 2008.

To accomplish our objectives, we interviewed NASS officials at the Headquarters level and the State field office level. We also reviewed relevant laws, regulations, procedures, and documents to obtain information regarding the agency's process for estimating county-level crop acreage, production, and yields.

Through research and interviews with agency officials, we determined that RMA and FSA are the two agencies that most frequently use NASS' county estimates to determine program benefits. We reviewed these agencies' program guidelines and procedures that use county-level estimates of crop yields to determine benefits. We also communicated with outside consultants and academia, for their opinions, perspective, and suggestions regarding the reliability of NASS' county yield data.

We analyzed RMA's relevant data sets in order to gain an understanding of its county-level farmer data. This data review encompassed crop years 2006 through 2008 for corn, soybeans, and wheat, and covered 11 major crop production States for corn and soybeans, and 11 major crop production States for wheat.³⁵

We conducted this audit in accordance with generally accepted government auditing standards. These standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

We assessed the reliability of NASS data by reviewing existing information about the data and the system that produced them, and interviewing agency officials knowledgeable about the data. We determined that the data were sufficiently reliable for the purposes of this report.

³⁵ The 11 States for corn and soybeans are: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, Ohio, South Dakota, and Wisconsin. The States that our data review covered for wheat were Colorado, Kansas, Missouri, Montana, North Dakota, Nebraska, Oklahoma, Oregon, South Dakota, Texas, and Washington.

Abbreviations

CAPS.....	County Agricultural Production Survey
CADW	Crop Acreage Data Warehouse
CES	County Estimates Survey
DICE	Database Integrated County Estimates
FSA	Farm Service Agency
GRP.....	Group Risk Plan
GRIP	Group Risk Income Protection
NASS	National Agricultural Statistics Service
OIG	Office of Inspector General
USDA.....	Department of Agriculture

**USDA'S
NATIONAL AGRICULTURAL
STATISTICS SERVICE
COMBINED RESPONSE TO
AUDIT REPORT**



United States Department of Agriculture
National Agricultural Statistics Service
Office of the Administrator



April 8, 2011

TO: Gil H. Harden
Assistant Inspector General for Audit

FROM: Dr. Cynthia Clark
Administrator
National Agricultural Statistics Service

A handwritten signature in black ink, appearing to read "Cynthia Clark".

SUBJECT: National Agricultural Statistics Service Establishment of Average Yields
(50601-15-KC)

Attached are the written comments pertaining to the official draft report. These comments combine the response of each of the three agencies identified in the recommendations, specifically, the National Agricultural Statistics Service (NASS), the Risk Management Agency (RMA), and the Farm Service Agency (FSA).

These agencies have historically worked well together and it is believed that the findings identified in your audit will only strengthen these working relationships.

If you have any questions or concerns, please contact Brad Summa at (202) 720-2707.

Recommendation No. 1

Recommendation number 1 states that NASS should, “Develop a strategy to use RMA acreage, production, and yield data to cross-check and supplement NASS’ process for generating county-level yield estimates.

The timing of reporting by producers to RMA has historically created a barrier for widespread use of these data by NASS. However, recent technological advancements make it worthwhile to again explore the use of RMA data and develop a strategy to utilize these data in the most appropriate fashion for NASS county-level yield estimates.

This exploration and plan on utilizing the data will be conducted by NASS during the 2011 production season. A targeted completion date for the review and resulting plan is April 1, 2012.

Recommendation No. 2

NASS and RMA have been working together over the past several years to determine optimal use by NASS of administrative data provided to RMA while implementing its insurance programs. This working relationship is documented in a Memorandum of Understanding (MOU) signed by both agencies in 2008.

Recommendation number 2 states that NASS and RMA should, “Establish a formal agreement giving NASS ongoing access to RMA data to the extent practicable given the systems of the agency’s reporting requirements.”

NASS and RMA have established a formal agreement giving NASS ongoing access to RMA data. This access is formalized in the attached document titled, “Interconnection Security Agreement between Risk Management Agency (RMA) and National Agricultural Statistics Service (NASS)” effective on December 14, 2010.

We request management decision for this recommendation.

Recommendation No. 3

Recommendation number 3 states that NASS and RMA should, “Implement the agreement by undertaking those tasks necessary by each agency to enable NASS full access to RMA data.”

NASS and RMA have included in the agreement “Interconnection Security Agreement between Risk Management Agency (RMA) and National Agricultural Statistics Service (NASS)” effective on December 14, 2010 the tasks necessary for each agency to perform in order to enable NASS full access to RMA data.

We request management decision for this recommendation.

Recommendation No. 4

Recommendation number 4 states that FSA should, “provide evidence that corrective actions for data mart’s deficiencies have been completed, including:

- a. Application controls to identify and prevent duplicate records;
- b. Accurate business logic incorporated to keep the system consistent with program changes;
- c. Division assigned responsibility for maintaining the data mart; and
- d. Data mart’s information validated on a regular basis.

FSA has completed development of the Crop Acreage Data Mart (CADW) changes required to implement the corrective measures recommended by OIG. The CADW is in the final certification testing phase prior to re-deployment. A team of FSA State and National Office (WDC) and National Agricultural Statistics Service (NASS) employees have been identified to test the data mart. Corrective measures will be verified by the certification testing effort. Once the verification is complete the Data Mart will be re-deployed to production.

In order to prevent duplicates from occurring, detail acreage reporting data will Primary keys were added to the CADW. This approach also allows detailed verification of data synchronized with the Mainframe acreage reporting data.

The AS/400, Mainframe and Data Mart files are monitored and synchronized daily. Control totals are contained within the daily transmission files and are balanced to the Mainframe data. Exception reports are created and monitored daily. These exceptions are researched. Upon completing the research, if required; a full file upload is triggered in order to synchronize the acreage data.

Subject matter experts (SME) from FSA’s Database management Office (DBMO) and the Production Adjustment and Risk Management Office (PARMO) are assigned to maintain the integrity of the CADW by reviewing daily changes and assessing the impacts.

Starting with Crop Year 2011, acreage reports are being captured in a centralized web base system. This simplifies validation as it eliminates synchronization of the data between the AS/400 and the Mainframe database that feeds the CADW.

The Productions, Emergencies, and Compliance Division (PECD) is now the Business owner for CADW. DBMO along with PARMO provide IT support for CADW.

As the CADW is currently not available to support NASS, WDC uses an EZTRIEVE mainframe (MF) program written by an FSA employee that extracts only the information (fields) requested by NASS to create a flat file that can be transferred for further processing to meet NASS’s request. Crop codes are totaled and validate independently for the United States. Any further processing that is done to the flat file (e.g. to county level summaries) is checked against the source file totals.

FSA provides NASS with complete and accurate acreage data from the mainframe data files. The data is received daily from the AS/400 which contains changes the county(s) have applied. Control totals are attached to the transmission files. The totals are checked and balanced daily. Exception reports are created and monitored on a daily basis. If required, a full file upload is performed in order to synchronize the acreage data.

FSA has added corrective action in order to provide NASS with complete and accurate acreage data from the NITC 578 Compliance Detail File.

Recommendation No. 5

NASS and FSA have been working closely together over the past several years to share data. This working relationship is documented in a Memorandum of Understanding (MOU) signed by both agencies in 2004.

Recommendation number 5 states that NASS and FSA should establish a formal agreement between the agencies that institutes the controls necessary to ensure the integrity of data needed with respect to its intended use.

The integrity of data used by NASS, or any other Federal statistical agency, is the responsibility of both the source of the information and the Federal statistical agency using the information. NASS and FSA will revise the existing MOU to address the controls necessary to ensure the integrity of the data with respect to its intended use. The MOU will be revised by September 30, 2011.

Informational copies of this report have been distributed to:

Administrator, NASS	
Attn: Agency Liaison Officer	****(3)
Administrator, RMA	
*****Cwp<Ci gpe{ "Nkukp"Qhhegt"	*****(4)
Administrator, FSA	
Attn: Agency Liaison Officer	****(2)
Government Accountability Office	****(1)
Office of Management and Budget	****(1)
Office of the Chief Financial Officer	
****Cwp<Director, Planning and Accountability Division "	(1)

To learn more about OIG, visit our website at
<https://www.usda.gov/oig/index.html>

How To Report Suspected Wrongdoing in USDA Programs

Fraud, Waste, and Abuse

In Washington, D.C.: (202) 690-1622

Outside D.C.: (800) 424-9121

TDD (Call Collect): (202) 690-1202

(Monday-Friday, 9:00 a.m. – 3:00 p.m. ET)

Bribes or Gratuities:

(202) 720-7257

(888) 620-4185

(24 hours a day)



The U.S. Department of Agriculture (USDA) prohibits discrimination in all its programs and activities on the basis of race, color, national origin, age, disability, and where applicable, sex, marital status, familial status, parental status, religion, sexual orientation, genetic information, political beliefs, reprisal, or because all or part of an individual's income is derived from any public assistance program. (Not all prohibited bases apply to all programs.) Persons with disabilities who require alternative means for communication of program information (Braille, large print, audiotape, etc.) should contact USDA's TARGET Center at (202) 720-2600 (voice and TDD). To file a discrimination complaint, write to USDA, Director, Office of Civil Rights, 1400 Independence Avenue, S.W., Washington, D.C. 20250-9410, or call (800) 795-3272 (voice) or (202) 720-6382 (TDD). USDA is an equal opportunity provider and employer.