



Неофициальный перевод – Статкомитет СНГ

Руководство по совершенствованию и использованию административных данных в сельскохозяйственной статистике



Февраль 2018

СОДЕРЖАНИЕ

СПИСОК СОКРАЩЕНИЙ.....	v
ВЫРАЖЕНИЕ ПРИЗНАТЕЛЬНОСТИ.....	vii
ПРЕДИСЛОВИЕ.....	viii
РЕЗЮМЕ.....	ix
ГЛАВА 1. РАЗРАБОТКА СИСТЕМЫ АДМИНИСТРАТИВНЫХ ДАННЫХ ДЛЯ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИКИ (ADSAS).....	12
1.1. ОПРЕДЕЛЕНИЯ АДМИНИСТРАТИВНЫХ ДАННЫХ И СИСТЕМЫ АДМИНИСТРАТИВНЫХ ДАННЫХ ДЛЯ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИКИ (ADSAS).....	12
1.2. ОПРЕДЕЛЕНИЕ ИСТОЧНИКОВ АДМИНИСТРАТИВНЫХ ДАННЫХ ПО СЕЛЬСКОМУ ХОЗЯЙСТВУ.....	15
1.3. ПОСТРОЕНИЕ СИСТЕМЫ АДМИНИСТРАТИВНЫХ ДАННЫХ ДЛЯ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИКИ (ADSAS).....	19
ГЛАВА 2. ОЦЕНКА КАЧЕСТВА СТРУКТУРЫ И ДАННЫХ В СИСТЕМЕ АДМИНИСТРАТИВНЫХ ДАННЫХ ДЛЯ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИКИ (ADSAS).....	22
2.1. ДИАГНОСТИКА СТРУКТУРЫ.....	22
2.2. ОЦЕНКА КАЧЕСТВА АДМИНИСТРАТИВНЫХ ДАННЫХ.....	23
ГЛАВА 3. ПОВЫШЕНИЕ КАЧЕСТВА АДМИНИСТРАТИВНЫХ ДАННЫХ ПО СЕЛЬСКОМУ ХОЗЯЙСТВУ.....	28
3.1. МЕТОДОЛОГИЧЕСКИЙ ИНСТРУМЕНТАРИЙ ДЛЯ РАССМОТРЕНИЯ ВОПРОСОВ КАЧЕСТВА.....	28
3.2. ЛЮДСКИЕ И ФИНАНСОВЫЕ РЕСУРСЫ.....	32
3.3. СТАНДАРТИЗАЦИЯ ПОНЯТИЙ В УЧРЕЖДЕНИЯХ, ЗАНИМАЮЩИХСЯ АДМИНИСТРАТИВНЫМИ ДАННЫМИ В СЕЛЬСКОМ ХОЗЯЙСТВЕ.....	34
ГЛАВА 4. УЛУЧШЕНИЕ ДОСТУПНОСТИ АДМИНИСТРАТИВНЫХ ДАННЫХ.....	36
4.1. ПРАВОВАЯ СИСТЕМА.....	36
4.2. НОРМАТИВНАЯ БАЗА.....	37
4.3. ВОПРОСЫ, СВЯЗАННЫЕ С КОНФИДЕНЦИАЛЬНОСТЬЮ И ВОСПРИЯТИЕМ ОБЩЕСТВЕННОСТИ.....	38
4.4. СОГЛАШЕНИЯ МЕЖДУ УЧРЕЖДЕНИЯМИ.....	38

ГЛАВА 5. ИСПОЛЬЗОВАНИЕ АДМИНИСТРАТИВНЫХ ДАННЫХ.....	43
5.1. МЕТОДОЛОГИЧЕСКИЕ И ТЕХНИЧЕСКИЕ ИНСТРУМЕНТЫ ДЛЯ УВЯЗКИ ЗАПИСЕЙ И ИНТЕГРАЦИИ ДАННЫХ	43
5.2. ВИДЫ ИСПОЛЬЗОВАНИЯ ДЛЯ ФОРМИРОВАНИЯ СТАТИСТИЧЕСКОГО ПРОДУКТА (КОСВЕННОЕ ИСПОЛЬЗОВАНИЕ).....	52
5.3. ИСПОЛЬЗОВАНИЕ В КАЧЕСТВЕ КОНЕЧНОГО ПРОДУКТА	61
ГЛАВА 6. ИНТЕГРИРОВАНИЕ АДМИНИСТРАТИВНЫХ ДАННЫХ В НАЦИОНАЛЬНЫЕ СИСТЕМЫ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИКИ.....	67
6.1. ИНТЕГРАЦИЯ ADSAS В SPARS.....	67
6.2. ДОЛГОСРОЧНАЯ ПЕРСПЕКТИВА ИНТЕГРАЦИИ АДМИНИСТРАТИВНЫХ ДАННЫХ В СТАТИСТИЧЕСКУЮ СИСТЕМУ	68
ЛИТЕРАТУРА.....	72
ПРИЛОЖЕНИЯ.....	81

ТАБЛИЦЫ

ТАБЛИЦА 1. ИСТОЧНИКИ АДМИНИСТРАТИВНЫХ ДАННЫХ ДЛЯ НЕКОТОРЫХ ЭЛЕМЕНТОВ ОСНОВНЫХ ДАННЫХ ГЛОБАЛЬНОЙ СТРАТЕГИИ	20
ТАБЛИЦА 2. АСПЕКТЫ ОЦЕНКИ КАЧЕСТВА.	24
ТАБЛИЦА 3. РЕЗУЛЬТАТЫ ОЦЕНКИ КАЧЕСТВА ARDS ОБЪЕДИНЕННОЙ РЕСПУБЛИКИ ТАНЗАНИЯ.....	26
ТАБЛИЦА 4. ТЕХНИЧЕСКИЕ ВОЗМОЖНОСТИ ПАКЕТОВ ДЛЯ УВЯЗКИ ЗАПИСЕЙ	49
ТАБЛИЦА 5. ОБЗОР УДОБСТВА ИСПОЛЬЗОВАНИЯ РАЗНЫХ ПАКЕТОВ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ	50
ТАБЛИЦА 6. ОСНОВНЫЕ СИЛЬНЫЕ И СЛАБЫЕ СТОРОНЫ ПАКЕТОВ ПРОГРАММ ДЛЯ УВЯЗКИ ЗАПИСЕЙ.....	51
ТАБЛИЦА 7. ЗНАЧЕНИЯ ДИСПЕРСИИ ДЛЯ ОЦЕНОК, ПОЛУЧЕННЫХ ИЗ НАЦИОНАЛЬНОЙ СЕЛЬСКОХОЗЯЙСТВЕННОЙ ПЕРЕПИСИ (NSA)	59
ТАБЛИЦА 8. АДМИНИСТРАТИВНОЕ ИСПОЛЬЗОВАНИЕ ADSAS: ИСПОЛЬЗОВАНИЕ ДЛЯ ФОРМИРОВАНИЯ СТАТИСТИКИ	62
ТАБЛИЦА 9. ПОГОЛОВЬЕ СКОТА В ВЕЛИКОБРИТАНИИ И СОЕДИНЕННОМ КОРОЛЕВСТВЕ НА 1 ИЮНЯ ПО ДАННЫМ ОБСЛЕДОВАНИЯ И ИЗ АДМИНИСТРАТИВНЫХ ИСТОЧНИКОВ: 2003 и 2006	63
ТАБЛИЦА 10. ИСПОЛЬЗОВАНИЕ ADSAS В КАЧЕСТВЕ ОКОНЧАТЕЛЬНОЙ СТАТИСТИКИ.....	66

РИСУНКИ

РИСУНОК 1. СИСТЕМА АДМИНИСТРАТИВНЫХ ДАННЫХ ДЛЯ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИКИ.....	14
РИСУНОК 2. ТИПИЧНАЯ АДМИНИСТРАТИВНАЯ СИСТЕМА ОТЧЕТНОСТИ В АЗИИ.....	19
РИСУНОК 3. ДВА АСПЕКТА ОЦЕНКИ КАЧЕСТВА АДМИНИСТРАТИВНЫХ ДАННЫХ.....	27
РИСУНОК 4. СХЕМА ОРГАНИЗАЦИИ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИКИ В СТАТИСТИЧЕСКОМ УПРАВЛЕНИИ КАНАДЫ.....	68
РИСУНОК 5. ПРОИЗВОДСТВЕННАЯ СИСТЕМА НА БАЗЕ РЕГИСТРОВ, СИСТЕМА РЕГИСТРОВ И ВЫБОРОЧНЫЕ ОБСЛЕДОВАНИЯ.	70

СПИСОК СОКРАЩЕНИЙ

ВВП	Валовый внутренний продукт	
МСХ США	Министерство сельского хозяйства США	
ООН	Организация Объединенных Наций	
НОСС	Национальная организация сельскохозяйственной статистики	
НСУ	Национальное статистическое управление	
ОЭСР	Организация экономического сотрудничества и развития	
ФАО	Продовольственная и сельскохозяйственная организация Объединенных наций	
ЦСУ	Центральное статистическое управление	
AAS	Ежегодное сельскохозяйственное обследование	<i>Annual Agricultural Survey</i>
ADSAS	Система административных данных для сельскохозяйственной статистики	<i>Administrative Data Systems for Agricultural Statistics</i>
ARDS	Система регулярных сельскохозяйственных данных	<i>Agricultural Routine Data System</i>
ARSA	Система административной отчетности для сельского хозяйства	<i>Administrative Reporting Systems for Agriculture</i>
ASDP	План развития сельскохозяйственного сектора	<i>Agricultural Sector Development Plan</i>
CAPI	Интервьюирование с помощью компьютера	<i>Computer-Assisted Personal Interview</i>
CV	Коэффициент вариации	<i>Coefficient of variation</i>
FSA	Агентство по обслуживанию фермеров при МСХ США	<i>Farm Services Agency (USDA)</i>
GPS	Глобальная система позиционирования	<i>Global Positioning System</i>
GSARS	Глобальная стратегия совершенствования сельскохозяйственной и сельской статистики	<i>Global Strategy to Improve Agricultural and Rural Statistics</i>
IACS	Интегрированная система администрирования и контроля	<i>Integrated Administrative and Control System</i>
ICS	Схема улучшения статистики сельскохозяйственных культур	<i>Improvement of Crop Statistics (India)</i>
IMDB	Интегрированная база метаданных	<i>Integrated Metadata Base</i>
LGMD	База данных мониторинга местных органов власти	<i>Local Government Management Database</i>

MAF	Департамент планирования и сотрудничества Министерства сельского и лесного хозяйства, Лаос	<i>Ministry of Agriculture and Forestry (Lao People's Democratic Republic)</i>
MAWF	Министерство сельского, водного и лесного хозяйства Намибии	<i>Ministry of Agriculture, Water and Forestry (Namibia)</i>
MMSE	Минимальная среднеквадратическая ошибка	<i>Minimum Mean Squared Error</i>
NASS	Национальная служба сельскохозяйственной статистики (США)	<i>National Agricultural Statistics Service (United States of America)</i>
NCA	Национальная сельскохозяйственная перепись	<i>National Census of Agriculture</i>
NRCS	Служба охраны природных ресурсов (США)	<i>National Resources Conservation Service (United States of America)</i>
NSA	Статистическое агентство Намибии	<i>Namibia Statistics Agency</i>
PPS	Вероятность отбора, пропорциональная размеру	<i>Probability Proportional to Size</i>
TAD	Традиционные административные данные	<i>Traditional Administrative Data</i>
TRS	Схема своевременной отчетности	<i>Timely Reporting Scheme</i>

ВЫРАЖЕНИЕ ПРИЗНАТЕЛЬНОСТИ

Этот документ основан на результатах проекта ADMIN Глобальной стратегии, осуществляемого Школой статистики и планирования / Колледжем бизнеса и управления Университета Макерере (Уганда) и Центром статистики и методологии обследований Университета Айовы. (США). Список ключевых экспертов из двух организаций, участвующих в проекте, и всех других участников можно найти в окончательном техническом отчете ADMIN: GSARS (2017).

Руководство было подготовлено Драмане Бако (*Dramane Bako*), международным консультантом-статистиком Глобальной стратегии совершенствования сельскохозяйственной и сельской статистики, под руководством Флавио Боллижера (*Flavio Bolliger*), координатора исследований Глобальной стратегии. Следующие эксперты Глобальной стратегии и ФАО рассмотрели документ и внесли ценный вклад: Наман Кейта (*Naman Keita*) - старший консультант по сельскохозяйственной статистике, Глобальная стратегия, Валери Бизье (*Valerie Bizier*) - координатор технической помощи и обучения по Глобальной стратегии, Кристоф Дюамель (*Christophe Duhamel*) – координатор, Глобальный офис Глобальной стратегии, Карола Фаби (*Carola Fabi*) - статистик, Статистический отдел ФАО, и Глэдис Кача (*Gladys Qacha*) - эксперт по сельскохозяйственной статистике, Глобальная стратегия.

Эта публикация была подготовлена при поддержке Целевого фонда Глобальной стратегии, финансируемого Департаментом международного развития Соединенного Королевства (DfID) и Фондом Билла и Мелинды Гейтс.

Документ был отформатирован Лорой Монополи (*Laura Monopoli*) под руководством Норы ДеФалько (*Norah DeFalco*).

Перевод на русский язык осуществлен Статкомитетом СНГ в рамках Гранта для развития сельскохозяйственной и сельской статистики в регионе Содружества Независимых Государств, финансируемого Всемирным банком.

ПРЕДИСЛОВИЕ

Глобальная стратегия совершенствования сельскохозяйственной и сельской статистики (далее - Глобальная стратегия или GSARS), принятая Статистической комиссией Организации Объединенных Наций в 2010 году, направлена на повышение качества и устойчивости статистических данных по сельскому хозяйству в развивающихся странах. Одним из ключевых компонентов Плана действий по реализации Глобальной стратегии является ее Программа исследований, которая обеспечивает поддержку исследований и разработки экономически эффективных методов, которые послужат основой для технических руководств, справочников и учебных материалов для использования консультантами, статистиками в странах и учебными центрами.

Глобальная стратегия реализует важное направление исследований по совершенствованию методологии использования административных данных в системе сельскохозяйственной статистики (ADMIN), одного из ключевых приоритетов исследовательской программы. В рамках ADMIN было разработано восемь технических отчетов. Они доступны по адресу <http://gsars.org/en/tag/administrativedata/>.

В рамках этой темы цель состоит в том, чтобы исследовать методы улучшения сбора, управления и использования административных данных для производства сельскохозяйственной статистики в развивающихся странах. Конечная цель этого документа - предоставить развивающимся странам оперативное руководство о том, как создать эффективную систему административных данных для сельскохозяйственной статистики (ADSAS), а также об улучшении, использовании и интеграции административных данных в национальной статистической системе. Концепция системы административных данных для сельскохозяйственной статистики (ADSAS) относится к *совокупности всех административных учреждений, которые производят административные сельскохозяйственные данные, которые могут использоваться для целей сельскохозяйственной статистики, и предоставляют их национальному учреждению, отвечающему за сельскохозяйственную статистику, для официального использования и публикации.*

Настоящее Руководство по совершенствованию и использованию административных данных в сельскохозяйственной статистике во многом совпадает с окончательным техническим отчетом ADMIN. Для разработки документа были использованы выводы, изложенные в других технических отчетах ADMIN, и дополнительные материалы из литературы.

РЕЗЮМЕ

Определение административных данных, принятое для этого документа, описывает эти данные как «информацию, собираемую главным образом для административных целей правительственными агентствами и другими организациями, обычно во время предоставления услуги или для целей регистрации, ведения учета или документирования операции»¹. Это определение охватывает два класса административных данных, обозначаемых соответственно термином «**Традиционные административные данные**» (**TAD**), которые включают в себя информацию, собираемую в рамках системы налогообложения, процессов регулирования (инспекции ферм), программ помощи фермерам (субсидии и страхование) и т.д., и «**Системы административной отчетности для сельского хозяйства**» (**ARSA**), которые относятся к данным, регулярно собираемым специалистами по распространению сельскохозяйственных знаний. Свойства этих двух типов данных различаются из-за различий между методами сбора данных.

Концепция системы административных данных для сельскохозяйственной статистики (ADSAS) относится к совокупности всех административных учреждений, производящих административные сельскохозяйственные данные, которые могут использоваться для целей сельскохозяйственной статистики, и предоставляющих их национальному учреждению, отвечающему за сельскохозяйственную статистику для использования и официальной публикации. Этим учреждением обычно является Национальное статистическое управление (НСУ), Центральное статистическое управление (ЦСУ) или отдел сельскохозяйственной статистики в министерстве, отвечающем за сельское хозяйство, которое в настоящем Руководстве будет обозначаться аббревиатурой НОСС (Национальная организация сельскохозяйственной статистики).

Административные данные имеют различные преимущества, как обсуждалось в Глобальной стратегии (GSARS, 2015a); их использование приводит к экономии затрат, снижению нагрузки на респондентов и повышению эффективности оценок на макроуровне и в статистике малых областей. Таким образом, создание эффективной системы административных данных для сельскохозяйственной статистики (ADSAS) может способствовать улучшению качества и доступности сельскохозяйственной статистики.

Разработка административных данных для сельскохозяйственной статистики (ADSAS) начинается с определения соответствующих административных источников, которые будут являться частью системы. Во многих развивающихся странах основные сельскохозяйственные административные данные (то есть данные по сельскохозяйственным культурам, животноводству, рыболовству и лесному хозяйству) собираются и управляются министерствами

См. www.adls.ac.uk/adls-resources/guidance/introduction/.

сельского хозяйства, животноводства, рыболовства и/или лесного хозяйства. Однако во многих других странах полугосударственные организации также представляют административные данные, особенно по коммерческим или товарным культурам. Агентства или организации частного сектора также часто собирают и обрабатывают различные виды административных данных, например, о ценах, маркетинге, ресурсах и т.д., особенно в соответствии с политикой реструктуризации, принятой во многих из этих стран. Эти агентства могут собирать данные и управлять ими без какого-либо прямого участия соответствующих НСУ или ЦСУ. Примеры программ правительственной помощи и регулирования, которые могут предоставлять административные данные, включают программы страхования урожая и субсидирования, информацию из ветеринарных инспекций и от кампаний по вакцинации скота. Такие источники данных предоставляют информацию о таких аспектах, как площадь сельскохозяйственных культур и поголовье скота. Примеры мониторинга и ведения учета включают регистрацию земли и кадастровые записи, а также обследования почвы - административные операции, в ходе которых полностью документируются типы почв определенного региона. Примерами организаций частного сектора, которые могут предоставлять административные данные, относящиеся к сельскому хозяйству, являются следующие: бюро лицензирования или регулирования, зерновые ассоциации, товарные ассоциации, кооперативы, фабрики, бойни и распространители сельскохозяйственного сырья.

На страновом уровне определение потенциальных источников административных данных, генерирующих информацию, связанную с сельскохозяйственным сектором, которая может быть использована для сельскохозяйственной статистики, может быть не простым. Практическим способом составления хорошей инвентаризации источников административных данных в сельском хозяйстве и привязки их к основным элементам данных может быть организация качественного обследования сельскохозяйственной информации, обрабатываемой или производимой в рамках административных структур, а затем проведение национального рабочего совещания по результатам обследования. Это совещание должно привести к составлению окончательного списка всех административных источников, которые могут внести вклад в сельскохозяйственную статистику.

Оценка качества системы административных данных для сельскохозяйственной статистики (ADSAS) заключается в структурной диагностике системы и оценке качества данных. Структурная диагностика заключается, главным образом, в анализе основных структурных проблем, с которыми сталкиваются учреждения, производящие сельскохозяйственные административные данные, таких как недостаточная согласованность концепций и определений или правовые и политические ограничения. Предлагается структура для оценки качества административных сельскохозяйственных данных. К параметрам качества относятся следующие: актуальность, точность, доступность, конфиденциальность и защита конфиденциальности, согласованность, своевременность, пунктуальность и сопоставимость.

Статистические агентства в развитых странах разработали механизмы для обеспечения качества данных, которые могут применяться в развивающихся странах. Взаимодействие с административными учреждениями и общественностью (например, в отношении обмена конфиденциальными административными данными) может помочь согласовать определения, требуемые для статистических целей, с определениями, используемыми административными органами, смягчить влияние административных изменений на применимость данных для статистических целей, а также подойти к решению проблем конфиденциальности и безопасности. В связи с важностью объединения нескольких источников данных для преодоления проблем недостаточного охвата, методам и процедурам увязки записей и наборов данных уделяется большое внимание в литературе об использовании административных данных для производства официальной статистики. Подходы, которые объединяют административные

базы данных с информацией, полученной в ходе обследований, могут уменьшить проблемы, связанные с ошибкой измерения, обеспечить согласованность определений в разных источниках и улучшить охват. Аудит и выборки из административных данных используются для проверки ошибок и оценки охвата. Принятие передовых методов контроля и обеспечения качества может помочь в управлении ошибками в административных данных.

Административные данные могут использоваться напрямую в качестве конечного статистического продукта или косвенно - при формировании статистического продукта. Прямое использование состоит в публикации непосредственно таблиц административных данных в качестве выходного продукта. Этот вид использования рекомендуется только при ряде условий, которые включают полный охват рассматриваемой целевой совокупности. Статистические управления часто косвенно используют административные данные для поддержки программ обследования или переписи. Административные данные могут играть роль на всех этапах процесса обследования: построение совокупности, сбор и оценка данных (калибровка, оценка малых областей и импутация). Использование административных данных в качестве части программы обследования или переписи особенно полезно, когда административные данные и данные обследования имеют разные сильные и слабые стороны. В обычной ситуации административные данные предоставляют по существу полную информацию (то есть перепись) о количестве, которая коррелирует с объектом, представляющим интерес для статистического управления, но отличается от него. Обследование, напротив, позволяет получить точную информацию для подвыборки.

Доступ к административным данным может иметь юридические и политические ограничения, которые могут быть введены по уважительной причине - например, для защиты конфиденциальности отдельных лиц в совокупности. Институциональные механизмы имеют решающее значение для обеспечения лучшего качества и доступа к административным данным. К ним относятся структуры, обеспеченные персоналом для увязки с другими источниками сельскохозяйственных данных. Важным аспектом создания эффективной системы административных данных для сельскохозяйственной статистики (ADSAS) является обеспечение того, чтобы деятельность по сбору, анализу и распространению данных координировалась и разделялась между различным учреждениям. Обычно для этого требуется заключение официального соглашения или меморандума о взаимопонимании с указанием обязательств участвующих учреждений. Подробный меморандум о взаимопонимании, объясняющий цели статистического управления и данные, необходимые административному органу для достижения этих целей, часто необходим для организации потока данных от административного органа в статистическое управление.

Учитывая потенциальное использование административных данных для улучшения доступности и качества сельскохозяйственной статистики, рекомендуется хорошая интеграция этих данных в систему сельскохозяйственной статистики. Эти данные должны учитываться в национальной системе или стратегии, связанной с производством сельскохозяйственной статистики. Долгосрочной перспективой является развитие системы сельскохозяйственной статистики на основе регистров путем разработки или совершенствования основных регистров в стране.

Глава 1

РАЗРАБОТКА СИСТЕМЫ АДМИНИСТРАТИВНЫХ ДАННЫХ ДЛЯ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИКИ (ADSAS)

В этой главе представлены основные аспекты, которые следует учитывать при разработке ADSAS на уровне стран. Она начинается с определения административных данных, принятых в настоящем Руководстве, после краткого обзора многих имеющихся определений (см. Раздел 1.2., GSARS (2015a) и раздел 1.1., GSARS (2015b)). Описаны типологии и общие источники административных данных в сельском хозяйстве, а также отмечены результаты полевых испытаний в ходе исследовательского проекта ADMIN, проведенного в рамках Глобальной стратегии для определения источников административных данных. Наконец, предлагается оперативное руководство по разработке ADSAS для отдельных стран.

1.1. ОПРЕДЕЛЕНИЯ АДМИНИСТРАТИВНЫХ ДАННЫХ И СИСТЕМЫ АДМИНИСТРАТИВНЫХ ДАННЫХ ДЛЯ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИКИ (ADSAS)

1.1.1. Административные данные

Административные данные могут быть получены из различных источников, включая государственную регистрацию, системы отчетности и частные организации. Это указывает на сложность проблемы определения административных данных.

Традиционно некоторые авторы определяют источники административных данных как массивы данных, которыми владеют другие органы государственного управления и которые собирают и используют в целях администрирования налогов, льгот или услуг. Возможно, наиболее полное из традиционных определений было предложено Гордоном Брэкстоуном (*Gordon Brackstone*) из Статистического управления Канады, который выделил следующие четыре отличительные черты административных данных (Brackstone, 1987):

- Агент, который предоставляет данные статистическому управлению, и единица, к которой относятся данные, различны (в отличие от большинства статистических обследований);
- Данные первоначально были собраны с определенной нестатистической целью, которая может повлиять на обработку источника;
- Цель состоит в полном охвате целевой совокупности;

- Контроль методов сбора и обработки административных данных возлагается на административный орган.

Организация Объединенных Наций (ООН; см. ООН, 2011 г.) рассматривает традиционное или «узкое» определение административного источника, которое включает только нестатистические источники в государственном секторе, тогда как более широкое определение будет также включать источники в частном секторе. Таким образом, в узком определении административные источники являются подмножеством вторичных источников, в то время как в более широком определении эти термины являются синонимами. Наблюдается все больше причин для предпочтения более широкого определения, включая: (а) усиление приватизации государственных функций; б) разработку статистики населения на основе регистров (ООН, 2011 год), увеличение производства данных частным сектором и их перепродажи с добавленной стоимостью; и (с) рост интереса пользователей к новым типам данных (Brackstone, 1987). В работе Сена (Sen, без даты) дается альтернативное определение: «административные данные отличаются от статистических данных, когда конкретная идентичность респондента или источника данных является центральной для использования данных».

В данном Руководстве для определения административных данных в контексте сельскохозяйственной статистики мы следуем GSARS (2017а): «*информация, собираемая в основном для административных (нестатистических) целей органами государственного управления и другими организациями, как правило, во время оказания услуги или для целей регистрации, ведения учета или документирования операции*»².

1.1.2. Системы административных данных для сельскохозяйственной статистики (ADSAS)

Принятое более широкое определение административных данных в контексте сельскохозяйственной и сельской статистики охватывает два больших класса этих данных. В первом классе данные представляют собой измерения для хорошо определенных фермерских хозяйств, возникающих в результате участия в программе. Примерами этого первого типа административных данных является информация, собранная с помощью программ, касающихся налогообложения и субсидий. При разработке официальной статистики в хорошо развитых статистических системах широко используется этот первый тип административных данных. Ко второму классу - в основном в развивающихся странах - принадлежат данные, которые предоставляет специалист по распространению сельскохозяйственных знаний (при оказании услуг по оказанию помощи фермерам), деревенский староста или другой сотрудник по сельскохозяйственным вопросам, который выносит решение на основе своих наблюдений и экспертного заключения и регулярно составляет отчет о сельскохозяйственных культурах или о животноводстве в его или ее области работы.

Два термина введены, чтобы различать эти два класса административных данных. Первый из них называется «**Традиционные административные данные**» (TAD). В контексте сельского хозяйства, TAD включает информацию, собранную при помощи системы налогообложения, процессов регулирования (то есть инспекций фермерских хозяйств), программ помощи фермерам (субсидии и страхование) и программ мониторинга (системы отслеживания скота). Термин «**Система административной отчетности для сельского хозяйства**» (ARSA)

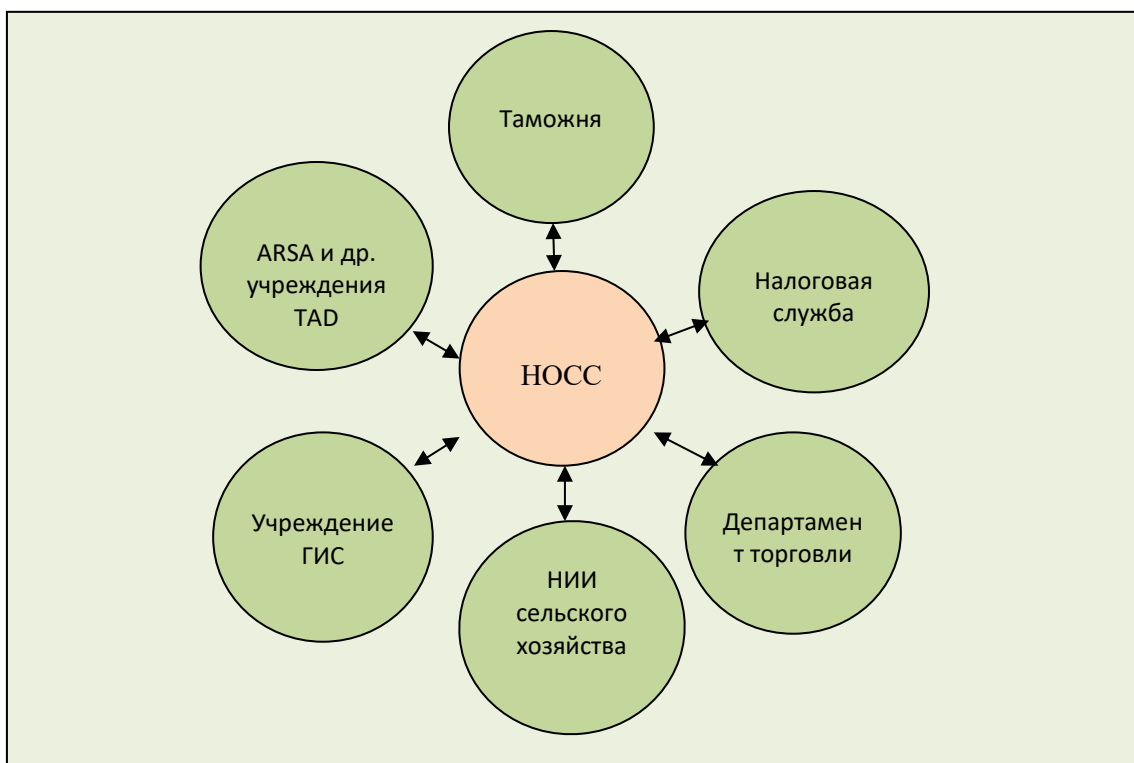
² См. www.adls.ac.uk/adls-resources/guidance/introduction/

описывает второй класс данных. Свойства этих двух типов данных различаются из-за различий в методах сбора данных. Для традиционных административных данных (ТАД) существует большой объем научно-методических работ. Однако в отношении системы административной отчетности для сельского хозяйства (ARSA) это не так; по этой причине этому второму классу административных данных в данном Руководстве уделяется больше внимания, поскольку во многих развивающихся странах в рамках существующих программ распространения знаний регулярно собирают информацию о сельском хозяйстве.

Система источников административных данных для сельскохозяйственной статистики (ADSAS) определяется как совокупность всех административных учреждений, производящих сельскохозяйственные административные данные, которые могут использоваться для целей сельскохозяйственной статистики, и предоставляющих их национальному учреждению, отвечающему за сельскохозяйственную статистику, для официального опубликования. В зависимости от статистической системы страны, этим учреждением может быть Национальное статистическое управление (НСУ), Центральное статистическое управление (ЦСУ) или департамент сельскохозяйственной статистики в министерстве, отвечающем за сельское хозяйство. Для упрощения в настоящем Руководстве соответствующее учреждение будет обозначено как Национальная организация сельскохозяйственной статистики (НОСС). Так же НСУ и ЦСУ будут просто обозначены аббревиатурой НСУ.

Система административных данных для сельскохозяйственной статистики (ADSAS) охватывает как традиционные административные данные (ТАД), так и систему административной отчетности для сельского хозяйства (ARSA). НОСС является ведущим учреждением системы административных данных для сельскохозяйственной статистики (ADSAS), собирающим данные из административных источников и использующим их для производства или улучшения сельскохозяйственной статистики в стране. В свою очередь, административные источники, участвующие в ADSAS, могут воспользоваться технической поддержкой или конкретной статистикой НОСС, соответствующей их потребностям. НОСС также может отвечать за разработку данных ARSA.

РИСУНОК 1. СИСТЕМА АДМИНИСТРАТИВНЫХ ДАННЫХ ДЛЯ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИКИ.



1.2. ОПРЕДЕЛЕНИЕ ИСТОЧНИКОВ АДМИНИСТРАТИВНЫХ ДАННЫХ ПО СЕЛЬСКОМУ ХОЗЯЙСТВУ

Во многих развивающихся странах основные сельскохозяйственные административные данные (то есть данные по сельскохозяйственным культурам, животноводству, рыболовству и лесному хозяйству) собираются и управляются министерствами сельского хозяйства, животноводства, рыболовства или лесного хозяйства. Однако во многих других странах полугосударственные организации также предоставляют административные данные, особенно по коммерческим или товарным культурам. Агентства или организации частного сектора также часто собирают и обрабатывают различные виды административных данных, например, о ценах, маркетинге, ресурсах и т.д., особенно в рамках политики реструктуризации, принятой во многих из этих стран. Эти агентства могут собирать данные и управлять ими без какого-либо прямого участия соответствующих НСУ.

В контексте сельского хозяйства административные источники включают (GSARS, 2015a):

- регулярную отчетность или отчеты сельскохозяйственных сотрудников или специалистов по распространению знаний (по различным аспектам, включая сельскохозяйственные культуры и скот);
- налоговые данные;
- записи о праве собственности на землю;
- информацию о государственных субсидиях;
- данные об импорте / экспорте;
- списки сельскохозяйственной продукции и сырья от производителей и дистрибьюторов;
- фермерские реестры и другие системы регистрации или лицензирования;
- данные по агротуризму;
- списки, которые ведут фермерские ассоциации;
- данные частного бизнеса;
- метеорологические данные;
- данные отслеживаемости, такие как данные по отслеживанию скота.

В этом разделе представлен обзор основных источников для двух категорий административных данных.

1.2.1. Традиционные административные данные (ТАД)

Основные источники традиционных административных данных рассмотрены ниже.

Информация о почве

Топографические карты и карты характеристик почвы часто поддерживаются в рамках административных процессов. Например, Служба охраны природных ресурсов (NRCS) является службой охраны почв Министерства сельского хозяйства США (USDA). NRCS поддерживает витрину данных о почве (Soil Data Mart) - базу данных о характеристиках почв в Соединенных Штатах Америки. Информация о характеристиках почвы из этих почвенных карт может быть использована для стратификации при обследованиях (см., например, Goebel, 2009) и в качестве вспомогательной информации при построении оценок.

Программы страхования урожая и субсидирования

Государственные программы помощи генерируют административные данные. Программы субсидий и страхования урожая могут собирать информацию, например, о площадях, засеянных определенными культурами (Carfagna and Carfagna, 2010). Доступ к административным базам данных государственных программ субсидирования и страхования требует установления хороших рабочих отношений с соответствующим административным органом (Prell et al., 2009). Эксперты-рецензенты подняли вопрос о том, что таких источников мало в развивающихся странах, хотя в некоторых докладах (Roberts, 2005; Clay, 2013) отмечается увеличение объема программ страхования, вводимых в развивающихся странах.

Некоторые примеры из опыта развитых стран:

Интегрированная система администрирования и контроля (IACS) Евростата: База данных, созданная для управления и контролирования выплат фермерам. IACS содержит информацию о посевных площадях на фермах, участвующих в программах субсидирования. Статистические управления Дании, Германии и Италии используют базу данных IACS для разных целей (FAO, 2010).

Агентство по обслуживанию фермеров при МСХ США (FSA) управляет несколькими сельскохозяйственными программами, включая субсидии и стимулы для сохранения сельскохозяйственных земель. Беклер (Beckler, 2013) описывает, как Национальная служба сельскохозяйственной статистики (NASS) использует данные FSA: «FSA является агентством МСХ США, и ему поручено администрировать различные программы сельскохозяйственной помощи и сохранения сельскохозяйственных земель, которые обеспечивают ценовую поддержку, помощь при бедствиях, кредиты и другие услуги сельхозпроизводителям. Сводный законопроект о фермерских хозяйствах США, который, как правило, обновляется каждые пять лет, предоставляет FSA полномочия по законотворчеству в рамках программ, которыми оно управляет. FSA собирает большой объем информации от сельскохозяйственных производителей при помощи различных форм заявок, необходимых для участия в программах. Некоторые из этих данных и данные географической информационной системы FSA используются NASS в качестве административных данных. NASS использует эти административные данные различными способами, включая: (1) построение и поддержку выборочной совокупности, (2) в качестве наземных достоверных данных для данных дистанционного зондирования и (3) в дополнение к данным, собранным в ходе переписей и обследований NASS».

Земельная регистрация и кадастровый учет

Кадастр состоит из записей, определяющих «протяженность, стоимость и собственность на землю» (Bins and Dale, 1995). Поддерживаемые несколькими развитыми странами и некоторыми развивающимися странами, кадастры используются для целей налогообложения, а также для предоставления точных описаний и непрерывного учета прав собственности на землю. Системы регистрации земли обычно имеют несколько общих характеристик с кадастрами и содержат обширную информацию о землепользовании, включая управление сельскохозяйственными культурами.

Например, в Индии существует децентрализованная статистическая система, в которой задачи распределяются между различными министерствами на национальном и региональном уровне. Система управления земельным налогом, администрируемая правительствами штатов, является источником административных данных, которые могут быть полезны при разработке сельскохозяйственной статистики (Sen, без даты; Goel, 2002). Этот ресурс состоит из информации о землепользовании и управлении сельскохозяйственными культурами, собранной учетчиками на уровне деревни. Примеры включают посевные площади, фруктовые сады, орошаемые площади и источники орошения (Goel, 2002). В системе регистрации земли, которая

охватывает 88 процентов посевных площадей (Goel, 2002), данные табулируются непосредственно из земельных регистров, а информация о регистрации используется в качестве совокупности для выборки для обследований урожайности и производства сельскохозяйственных культур (Goel, 2002; Sen, без даты). Целью системы своевременной отчетности, в рамках которой старосты деревень собирают данные для 20-процентной подвыборки вместо данных о всех посевных площадях, является ускорение сбора данных (Республика Индия, 2013).

Налоговые данные

Налоговые данные часто предоставляют информацию о расходах фермы. Данные о налогообложении уже давно используются в статистических процессах (Nordbotten, 2008), предоставляя информацию о доходах отдельных лиц и домохозяйств, типах и размерах бизнеса и изменениях, таких как миграция, а также информацию о начале или завершении коммерческих операций в те годы, когда не проводятся переписи. Различные роли налоговых данных в формировании краткосрочной бизнес-статистики описаны ОЭСР (2015). Статистические управления в Европе, Австралии, Канаде и США широко используют данные системы налогообложения при производстве сельскохозяйственной статистики.

Программы государственного регулирования и мониторинга

Программы государственного регулирования и мониторинга, будь то добровольные или обязательные, предоставляют значительный объем административных данных. Регулирующая деятельность включает мониторинг производственных процессов, финансовых учреждений и практики страхования, а полученные административные данные используются статистическими управлениями различными способами.

Данные о регулировании и мониторинге сельскохозяйственного производства играют важную роль в сельскохозяйственной статистике: в некоторых странах землевладельцы обязаны регистрировать свои земли, и информация также может быть получена из проверок безопасности пищевых продуктов на ферме и состоянии здоровья и записей о вакцинации. Такие системы, как системы отслеживания скота, которые отслеживают рождения, смерти и перемещение зарегистрированного скота, становятся все более важными в качестве источников административных данных. Некоторые примеры приведены ниже:

- Примером базы данных, созданной для регулирования отрасли, которая также может использоваться для сельскохозяйственной статистики, является бельгийская база данных SANITEL. SANITEL - это реляционная база данных, которая была создана для регулирования отрасли крупного рогатого скота и свиноводства и содержит постоянную инвентаризацию животных в Бельгии. Она предоставляет полный перечень поголовья и перемещений крупного рогатого скота и свиней, а также содержит информацию о состоянии здоровья животных и обнаружении антибиотиков, гормонов или загрязняющих веществ. Базу данных ведет Центральная ассоциация по охране здоровья животных, а не статистическое управление. Информацию в базе данных получают в результате регулирующей деятельности: каждый свиновод должен оформить сертификат о состоянии здоровья, где указывается размер его хозяйства. Впоследствии, примерно каждые три или четыре месяца, его посещает утвержденный ветеринар, чтобы он мог объявить тип и количество животных в действительности (European Communities, 2003). С 2002 года Бельгия сократила число обследований свиней с четырех до двух, предполагая заменить данные обследования информацией из SANITEL для составления своих прогнозов валового внутреннего производства (European Communities, 2003);
- Министерство сельского, водного и лесного хозяйства Намибии (MAWF) собирает обширные данные о животноводстве в рамках финансируемых правительством программ

вакцинации и мониторинга. Ежегодная перепись сельскохозяйственных животных, проводимая в рамках ежегодной кампании вакцинации, приводит к учету сельскохозяйственных животных в коммунальных и коммерческих сельскохозяйственных операциях. Кроме того, Намибия имеет систему отслеживания скота, которая позволяет отслеживать рождения и смерти, а также передвижение крупного рогатого скота. Интересно, что Намибия в настоящее время является единственной африканской страной с всеобъемлющей системой отслеживания скота.

Частный сектор и ассоциации

Частные организации, занимающиеся деятельностью, связанной с сельским хозяйством, такие как лицензионные или регулирующие бюро, зерновые ассоциации, товарные ассоциации, кооперативы, фабрики, бойни, дистрибьюторы сельскохозяйственной продукции и работники по распространению сельскохозяйственных знаний, связанные с университетами, регулярно собирают сельскохозяйственную информацию, которая может использоваться в официальной статистике. (USDA, 2011).

Мясной Совет и Сельскохозяйственный Союз Намибии являются двумя административными учреждениями в рамках Министерства сельского, водного и лесного хозяйства (MAWF). Они предоставляют Статистическому агентству Намибии (NSA) данные, необходимые для составления ежемесячного отчета по животноводству, который содержит информацию о количестве продаваемых сельскохозяйственных животных и индексах, измеряющих изменения по сравнению с базисным 2010 годом (NSA, 2015).

Кейта и Чин (Keita and Chin, 2013) ссылаются на исследование, проведенное в Кабо-Верде, в котором информация от частных организаций была очень важной, поскольку отсутствовали согласованные данные обследований или переписей. Они объясняют, что Кабо-Верде «является островной страной с орошаемым земледелием и товарными культурами, сосредоточенными в ограниченном числе хорошо известных зон», и что местная культура сельскохозяйственного производства способствует созданию системы фермерских организаций и кооперативов для определенных товарных культур.

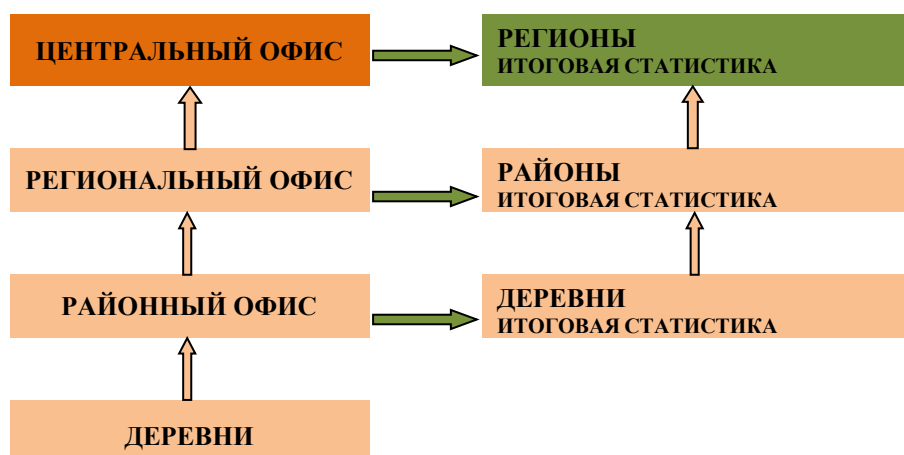
1.2.2. Системы административной отчетности для сельского хозяйства

Эксперты и специалисты в конкретных областях, которые регулярно участвуют в процессах сельскохозяйственного производства и в исследованиях, естественно, приобретают значительные экспертные знания. Субъективные оценки докладчиков-экспертов представляют собой полезную информацию для статистических управлений многих стран (Keita и Chin, 2013; Galmes, 2013; Namer, 2013). В число таких докладчиков входят эксперты, работающие в агробизнесе, университетских исследовательских лабораториях и административных агентствах (Namer, 2013), которые часто обладают экспертными знаниями в определенной области.

Системы отчетов экспертов особенно интересны для данного Руководства, поскольку они преобладают в развивающихся странах. В Африке системы сельскохозяйственной отчетности, созданные министерствами сельского хозяйства, могут предоставлять еженедельные, ежемесячные, полугодовые или годовые отчеты о посадках, производстве, состоянии урожая и погоде. Сбор рутинных административных сельскохозяйственных данных часто осуществляется через соответствующие министерства сельского хозяйства, животноводства, лесного и рыбного хозяйства на регулярной (еженедельной, ежемесячной или ежегодной) основе. Часто эти системы административной отчетности даже предоставляют данные о самых маленьких административных единицах, таких как районы или деревни. Примеры рутинных систем отчетности в развивающихся странах приведены ниже:

- Система регулярных сельскохозяйственных данных (ARDS) является основным источником информации о сельском хозяйстве в Объединенной Республике Танзания. ARDS была разработана в рамках Плана развития сельскохозяйственного сектора (ASDP) в ходе консультаций с несколькими регионами и районами с целью удовлетворения потребностей в данных для мониторинга и оценки самой ASDP. Стандартизированные данные собираются на уровне села и агрегируются на уровне участков, районов или регионов и, наконец, страны;
- Министерство сельского, водного и лесного хозяйства (MAWF) Намибии использует вопросник, который называется Список вопросов для оценки урожая (также известен как Список вопросов о производстве зерновых культур). Список вопросов позволяет собрать различную качественную и количественную информацию: погодные условия и состояние посевов, процент площади, засеянной зерновыми, оценка производства и пр.;
- Статистическая система Лаосской Народно-Демократической Республики децентрализована и включает в себя несколько учреждений, каждое из которых имеет свое конкретное назначение. Ежегодник сельскохозяйственной статистики является ежегодным изданием Департамента планирования и сотрудничества Министерства сельского и лесного хозяйства (MAF); в нем представлены сельскохозяйственные данные, полученные на основе административных отчетов. Большинство данных о производстве сельскохозяйственных культур и другие ряды сельскохозяйственных данных поступают из административных отчетов, в которых правительственный сельскохозяйственный персонал оценивает производство сельскохозяйственных культур путем наблюдения за урожаем и опроса ключевых информаторов (как правило, фермеров и старост деревень) в своих населенных пунктах.

РИСУНОК 2. ТИПИЧНАЯ АДМИНИСТРАТИВНАЯ СИСТЕМА ОТЧЕТНОСТИ В АЗИИ.



1.3. ПОСТРОЕНИЕ СИСТЕМЫ АДМИНИСТРАТИВНЫХ ДАННЫХ ДЛЯ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИКИ (ADSAS)

На уровне стран выявление потенциальных источников административных данных, предоставляющих информацию, связанную с сельскохозяйственным сектором и пригодных для использования в сельскохозяйственной статистике, может быть не простым. В этой главе приведены некоторые примеры источников с практическими описаниями для некоторых стран.

Однако этот список, конечно, не является исчерпывающим для многих стран, и не каждая страна будет иметь доступ к этим источникам. После инвентаризации соответствующих источников административных данных по сельскому хозяйству следующим важным шагом является увязка источников с основными элементами данных и другими важными национальными требованиями к данным.

Обсуждаемые выше источники охватывают многие элементы основных данных Глобальной стратегии. Данные о субсидиях предоставляют информацию о посевных площадях, налоговые данные - о затратах на ферме, отчетность боен и данные о вакцинации дают информацию для прогнозирования и оценки запасов сельскохозяйственных животных, а данные от дистрибьюторов предоставляют информацию о производстве молочной продукции (см. Таблицу 1 ниже).

ТАБЛИЦА 1. ИСТОЧНИКИ АДМИНИСТРАТИВНЫХ ДАННЫХ ДЛЯ НЕКОТОРЫХ ЭЛЕМЕНТОВ ОСНОВНЫХ ДАННЫХ ГЛОБАЛЬНОЙ СТРАТЕГИИ

Элементы основных данных	Тип административных данных	Пример
Сельскохозяйственные культуры <ul style="list-style-type: none"> • посевная площадь, площадь сбора урожая, урожай, хранение урожая, труд, цены • кукуруза, ячмень, пшеница, сорго, рис, хлопок 	Фермерские субсидии	Система IACS содержит данные о площадях под сельскохозяйственными культурами для культур, включенных в субсидии
	Ассоциации производителей	Ассоциация зерновых фермеров Онтарио предоставляет информацию о ценах
Животноводство <ul style="list-style-type: none"> • крупный рогатый скот, овцы, свиньи, козы, птица • инвентаризация, рождения, цены производства 	Регулирование здоровья животных	База данных SANITEL в Бельгии содержит данные организаций по регулированию здоровья животных и обследований добавок
	Системы отслеживания крупного рогатого скота	Системы отслеживания скота поставляют данные в Регистр крупного рогатого скота ЕС
Лесное хозяйство <ul style="list-style-type: none"> • Площадь лесов и лесопокрытых земель, количество вырубленного леса, цены 	Площадь лесопокрытых земель	Соединенное Королевство: Данные Комиссии по лесному хозяйству дополняют статистические обследования в оценке площади лесов и цен лесопокрытых земель в стране
Покров земли <ul style="list-style-type: none"> • Классификация покрова земли • категории: под сельскохозяйственными культурами; заболоченные земли, земли под лугами и пастбищами; 	Регистрация земель и кадастр	Индия: система регистрации земли в стране поддерживает оценки площадей в различных категориях земного покрова
Рыболовство/рыбоводство <ul style="list-style-type: none"> • Список крупных судов 	Административные записи о рыболовецких лодках/судах	Европейский союз: большая часть данных по рыболовству/рыбоводству, собранных Евростатом, основана на административных данных, таких как национальные реестры рыболовецких судов

Как указывалось выше, НСУ является ведущим учреждением ADSAS. Соответственно, это учреждение должно играть центральную роль в разработке ADSAS.

Практический способ создания хорошей инвентаризации источников административных данных в сельском хозяйстве и привязки их к основным элементам данных может предусматривать следующие шаги:

- организовать качественное обследование сельскохозяйственной информации, обрабатываемой и / или производимой в административных структурах в государственном секторе (государственное управление), гражданском обществе (НПО, ассоциации фермеров и т.д.) и в частном секторе; а потом
- организовать национальный семинар для иллюстрации результатов обследования с участием представителей потенциальных источников административных данных и их потенциальных пользователей (НСУ, министерства сельского хозяйства и т.д.). Такой семинар позволит уточнить результаты инвентаризационного обследования и оценить его сильные и слабые стороны и пригодность для использования в сельскохозяйственной статистике в рамках интегрированной и экономически эффективной системы сельскохозяйственной статистики. Этот семинар должен привести к разработке окончательного списка всех административных источников, которые могут внести вклад в сельскохозяйственную статистику.

РЕЗЮМЕ

В этой главе представлены определения, принятые в настоящем Руководстве для понятий административных данных и системы административных данных для сельскохозяйственной статистики (ADSAS). Приведены примеры обычных источников административных сельскохозяйственных данных для обеих категорий административных данных. Операционные этапы проектирования ADSAS предлагаются под руководством НОСС, начиная с инвентаризации источников административных данных и их привязки к ключевым данным. Это можно сделать с помощью качественного обследования, за которым следует национальный семинар. Следующий шаг будет состоять в оценке качества данных, полученных на этих этапах.

ГЛАВА 2

ОЦЕНКА КАЧЕСТВА СТРУКТУРЫ И ДАННЫХ В СИСТЕМЕ АДМИНИСТРАТИВНЫХ ДАННЫХ ДЛЯ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИКИ (ADSAS)

Оценка качества является важнейшим шагом для выявления проблем качества, влияющих на ADSAS с точки зрения улучшения сельскохозяйственных административных данных страны. Результаты оценки должны быть распространены среди пользователей данных системы административной отчетности для сельского хозяйства (ARSA) и традиционных административных данных (TAD), а также среди сборщиков данных. Это один из способов повышения осведомленности различных пользователей, в том числе государственных учреждений, о важности данных хорошего качества для информирования разработчиков политики и мониторинга политики (Maligalig, 2017). Следовательно, в рамках правительств может быть оказана поддержка для перехода или объединения результатов обследования и данных из ARSA и TAD для достижения конечной цели улучшения качества сельскохозяйственной и сельской статистики в стране.\

2.1. ДИАГНОСТИКА СТРУКТУРЫ

Диагностика структуры состоит главным образом в анализе основных структурных проблем, с которыми сталкиваются производители сельскохозяйственных административных данных..

Согласованы ли понятия и определения между источниками административных данных в сельском хозяйстве?

Как обсуждалось в главе 1, данные создаются различными учреждениями. Часто используются разные понятия и определения, что может привести к тому, что данные по одному и тому же элементу будут разными.

В связи с этим представляется необходимым изучить все концепции и определения, принятые целевыми учреждениями, производящими сельскохозяйственные административные данные, для выявления и понимания потенциальных различий и изучения путей стандартизации

Имеет ли персонал, вовлеченный в производство административных данных, достаточную квалификацию?

Во многих министерствах, департаментах и агентствах, отвечающих за сбор данных и управление ими, есть сотрудники в штаб-квартире и на местах (персонал по распространению знаний и сельские старосты или даже счетчики). Однако во многих развивающихся странах

высококвалифицированный персонал часто не может быть сохранен из-за плохих условий труда и недостаточных стимулов; это приводит к нехватке квалифицированного персонала. Все эти факторы способствуют низкому качеству генерируемых данных. Вторым недостатком является то, что руководство сотрудниками на местах часто осуществляется плохо.

Таким образом, оценка квалификации человеческих ресурсов, задействованных в работе с сельскохозяйственными административными данными, будет полезна для повышения качества данных.

Как контролируются сбор и обработка данных?

В отличие от статистических обследований и переписей административные данные собираются для целей, которые отличаются от целей статистических операций. Данные могут собираться отдельными лицами в неконтролируемых условиях без применения строгих протоколов, которые обязательны для процессов сбора данных тщательно проводимых обследований и переписей.

Обычно метод, используемый для сбора административных данных, в значительной степени находится вне контроля статистического агентства. Например, налоговые формы обычно заполняются отдельными налогоплательщиками. Во многих развивающихся странах доходы от сельскохозяйственной деятельности записываются специалистами по распространению сельскохозяйственных знаний или даже старостами деревень. Статистические отделы таможи или министерства сельского хозяйства, соответственно, часто не будут иметь никакого контроля. Эти формы сбора данных не стандартизированы и могут привести к ошибкам и несоответствиям в отчетности (ООН, 2011). В некоторых случаях искажения могут возникать из-за программных сбоев (Brackstone, 1987; Carfagna and Carfagna, 2010).

Рекомендуется проведение оценки процессов сбора и ввода данных. В частности, следует проверить наличие протоколов, связанных с этими процессами, и механизмов контроля качества в административных источниках.

Существуют ли правовые и политические ограничения?

Доступ к административным данным может быть ограничен юридическими и политическими ограничениями, которые могут иметь место небезосновательно (Brackstone, 1987). Одним из аспектов этих ограничений является конфиденциальность. НОСС обязано сохранять конфиденциальность статистических данных. Однако сложность этих требований возрастает, когда рассматриваются административные данные, которые собираются и хранятся другими учреждениями.

Для каждого производителя сельскохозяйственных административных данных важно определить все правовые и политические ограничения, касающиеся доступа к целевым данным, их использования и публикации.

2.2. ОЦЕНКА КАЧЕСТВА АДМИНИСТРАТИВНЫХ ДАННЫХ

Предложена система оценки качества сельскохозяйственных административных данных. Особое внимание уделяется использованию проверочных выборочных обследований.

2.2.1. Структура оценки качества

Глобальная стратегия (GSARS, 2016a) предлагает структуру оценки качества административной сельскохозяйственной статистики. Эта структура может использоваться для оценки качества как системы административной отчетности для сельского хозяйства (ARSA), так и традиционных

административных данных (TAD). В рамках исследовательского проекта ADMIN Глобальной стратегии она использовалась в ходе внутреннего тестирования для оценки качества Системы регулярных сельскохозяйственных данных (ARDS) в Объединенной Республике Танзания, которая является системой административной отчетности для сельского хозяйства (ARSA). Таблица 2 ниже иллюстрирует аспекты качества, которые считались актуальными для административных данных, и способы измерения и оценки каждого аспекта. Там, где это было возможно, предпринимались попытки сделать эти меры количественными.

ТАБЛИЦА 2. АСПЕКТЫ ОЦЕНКИ КАЧЕСТВА.

Аспект	Описание	Метод оценки
Актуальность	Степень, в которой доступная статистика соответствует потребностям нынешних и потенциальных пользователей. Это измерение также охватывает методологическую обоснованность и степень, в которой используемые концепции отражают потребности пользователей.	<p>Убедиться в интерпретируемости данных. Легко ли пользователи понимают данные?</p> <p>Провести обсуждения в фокус-группах с заинтересованными сторонами. Какие административные данные понятны или полезны?</p> <p>Существуют ли четкие определения понятий, целевых совокупностей, переменных и терминологии, а также информации, описывающей ограничения данных?</p> <p>Актуальны ли данные для объекта измерения?</p> <p>Каждый набор административных данных должен сопровождаться метаданными о его содержании, чтобы пользователи могли оценить пригодность набора данных для своих целей.</p>
Точность и надежность	<p>Близость статистических оценок к истинным значениям.</p> <p>Данные должны правильно оценивать или описывать количества или измеряемые характеристики.</p> <p>Точность также может быть описана в терминах основных источников ошибок, которые могут вызвать неточность (таких как охват, выборка, процент ответов и неответов).</p> <p>Данные должны адекватно представлять всю совокупность (полный охват) и соответствующие подгруппы (дезагрегация).</p>	<p>Данные должны производиться в соответствии с соответствующими стандартами, классификациями и практикой.</p> <p>Если производится отбор, необходимо выяснить, соответствует ли он стандартной схеме выборки.</p> <p>Следует определить процент отвечающих критериям респондентов, которые не были включены в учетные записи.</p> <p>Данные, собранные с использованием различных способов сбора, должны сравниваться (предлагаемые эксперименты также проинформируют об уровне точности).</p> <p>Коэффициенты вариации (CV) должны быть вычислены, когда это возможно.</p> <p>Административные данные следует сравнивать с данными обследований или даже переписей всякий раз, когда такие данные становятся доступными ³.</p>

³ Trant, 2010.

Аспект	Описание	Метод оценки
Доступность, конфиденциальность и защита частной информации	<p>Доступность следует рассматривать с точки зрения доступности для конечных пользователей данных.</p> <p>Термин может также охватывать спрос или фактический спрос на данные.</p>	<p>Данные могут быть легко найдены и доступны в различных форматах распространения (бумага, файлы, CD-ROM, Интернет и т. д.).</p> <p>Доступны метаданные, которые объясняют переменные и единицы измерения.</p> <p>Сводные отчеты и микроданные имеются и могут быть доступны для исследовательских целей.</p> <p>Количество пользователей данных и частота использования ими данных известны.</p> <p>Имеется четкая информация о том, где можно получить информацию, как ее заказать и каков срок ее доставки. Ценовая политика понятна и созданы удобные условия для рынка (авторские права и т.д.).</p>
Обоснованность и согласованность	<p>Данные из разных источников - и, в частности, из статистических обследований различной природы или частоты - могут быть не полностью согласованными, поскольку они могут основываться на различных подходах, классификациях или методологиях.</p> <p>Следовательно, такие данные могут не передавать полностью связанное сообщение пользователям (например, пользователи могут быть сбиты с толку, если два разных измерения одной и той же переменной публикуются с разными значениями).</p>	<p>Сравнение данных и связи между административными данными и данными обследований могут рассматриваться как возможный критерий для оценки согласованности административных данных в системе сельскохозяйственной статистики, так же как и анализ рядов административных данных.</p> <p>Сравнить данные из административных источников с данными переписей и обследований.</p> <p>Сравнить данные с другими внешними источниками, например, данными со спутниковых снимков. Использование спутниковых данных для оценки статистики землепользования стало более распространенным с увеличением доступности недорогих спутниковых изображений (Maligalig, 2017).</p> <p>Сравните подходы, классификации и методологии, используемые при сборе и анализе административных данных, с теми, которые используются при проведении переписей и обследований.</p>
Своевременность и пунктуальность	<p>Этот аспект касается непрерывного и последовательного распространения информации среди заинтересованных сторон, когда это необходимо.</p>	<p>Измерить промежуток времени между предоставляемыми данными (дата публикации) и событием или явлением, которое они описывают.</p> <p>Определить временную задержку между датой фактического выпуска данных (дата публикации) и целевой датой выпуска (часто предварительно объявленной).</p> <p>Как часто обновляются данные - как часто и в какие моменты времени?</p> <p>Убедиться, что отчетный период четко указан, чтобы можно было внести соответствующие корректировки, если административные данные должны быть интегрированы с обследованиями. Данные должны быть доступны пользователям при необходимости.</p>
Сопоставимость	<p>Этот аспект ориентирован на правильность сравнений между административными источниками и данными переписей и обследований, а также на правильность сопоставлений во времени и пространстве в рамках одного источника.</p>	<p>Одни и те же характеристики данных следует сравнивать между различными административными источниками и с данными переписей и обследований.</p> <p>Это также должно происходить в рамках одного источника, во времени и пространстве.</p>

Результаты оценки качества системы регулярных сельскохозяйственных данных (ARDS) Объединенной Республики Танзания представлены в Таблице 3.

ТАБЛИЦА 3. РЕЗУЛЬТАТЫ ОЦЕНКИ КАЧЕСТВА ARDS ОБЪЕДИНЕННОЙ РЕСПУБЛИКИ ТАНЗАНИЯ

Аспект	Оценка на основе ARDS в Объединенной Республике Танзания
Актуальность	<p><i>Ограничения:</i></p> <ul style="list-style-type: none"> • не все деревни охвачены; • деревенские дневники не стандартизированы; • технические заинтересованные стороны на национальном уровне - специалисты в области растениеводства и животноводства - не очень заинтересованы в использовании данных, полученных с помощью ARDS. Специалисты по растениеводству и животноводству продолжают собирать собственные данные для планирования и разработки политики.
Точность и надежность	<ul style="list-style-type: none"> • Гармонизированные форматы отчетности для всех регионов. • Гармонизированная частота отчетности. • Дневники для деревень не стандартизованы. • Используется случайный отбор. • Данные о производстве в значительной степени представляют собой оценки, а не фактические данные о производстве в домохозяйствах. Специалисты по распространению сельскохозяйственных знаний в деревнях должны заниматься измерением, но они не всегда это делают. Данные часто представляют собой прикидки специалистов по распространению сельскохозяйственных знаний. • Недостаточный надзор. • Отсутствует проверка данных. • Региональные менеджеры утверждают данные, введенные на уровне района, проверяя правильность информации. • Очень немногие районы в настоящее время отчитываются (18 процентов) из-за сбоя программного обеспечения, с одной стороны, и с другой, из-за недостатка знаний программного обеспечения, используемого для сбора и разработки данных, у персонала на более низких уровнях. Недостаточный персонал на местах и плохие условия труда, такие как отсутствие транспорта, также способствуют низкому уровню отчетности. • Данные составляются на уровне деревни, участка, района, региона и страны.
Доступность, конфиденциальность и защита частной информации	<ul style="list-style-type: none"> • Данные доступны на веб-сайте и на CD-ROM; тем не менее, доступность по-прежнему является проблемой. • Веб-сайт фиксирует количество людей, получающих доступ к веб-сайту для получения информации. • Отсутствие метаданных.
Обоснованность и согласованность	<p>Отсутствие сопоставимости данных ARDS и данных переписей и обследований из-за различий в инструментах сбора данных, сроках сбора данных, методологии, учетных периодах и т.д.</p>
Своевременность и пунктуальность	<ul style="list-style-type: none"> • Данные составляются ежемесячно, ежеквартально и ежегодно на уровне деревни, избирательного участка и района соответственно; и составляется ежеквартально и ежегодно на региональном и национальном уровнях. • Отчеты часто либо опаздывают, либо отправляются нерегулярно, что приводит к низкому проценту получения (18 процентов).
Сопоставимость	<p>Отсутствие сопоставимости данных ARDS с данными переписей и обследований из-за различий: в инструментах сбора данных, сроках сбора данных, методологии, учетных периодах и т.д.</p>

2.2.2. Проверка качества административных данных при помощи выборочного обследования

Проверка качества данных ARSA и TAD может быть выполнена посредством проведения выборочного обследования, чтобы оценить их точность и согласованность.

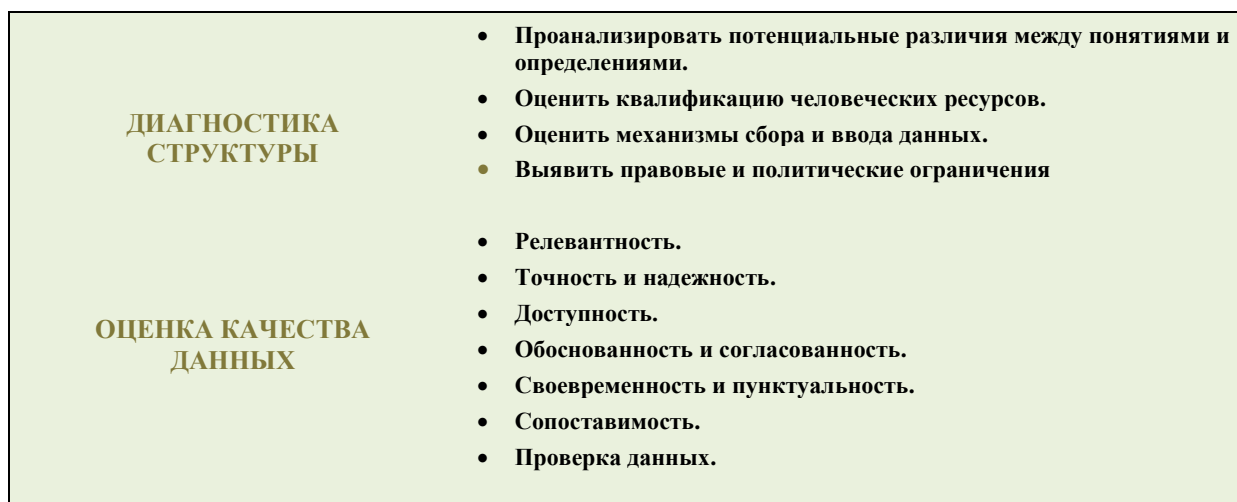
Что касается TAD, то для выявления ошибочных случаев может быть проведен опрос или проверка, в которых будет отмечены необычные записи. Несколько иной дизайн требуется, если целью является оценка степени ошибки в административном источнике (GSARS, 2017).

- Для ARSA параллельное выборочное обследование может использоваться для проверки качества отчетных данных на различных уровнях: деревня, провинция, страна и т.д. Maligalig (2017) предлагает процедуру выборки для проверки качества данных ARSA.
- Деревни могут быть отобраны с вероятностью, так что данные из всех выбранных деревень могут быть агрегированы, например, для получения оценок на районном уровне.
- Сотрудники отдела сельскохозяйственной статистики в министерстве сельского хозяйства могут обучать сборщиков данных или респондентов из отобранных деревень стандартным концепциям и определениям, а также важности предоставления точных и своевременных данных для органов, занимающихся выработкой политики и мониторингом.
- Персонал отдела сельскохозяйственной статистики может получить копию вопросника или формы на уровне деревни и объединить их в файл данных, чтобы можно было получить оценки на уровне района, провинции и страны, используя обратные вероятности выбора деревни в качестве весов.
- Эти оценки затем можно сравнить с результатами ARS. Затем сотрудники отдела сельскохозяйственной статистики должны внимательно изучить области с большими расхождениями.

РЕЗЮМЕ

В этой главе обсуждается оценка качества административных данных, которая имеет два измерения: диагностика структуры и оценка качества данных, производимых административными органами. Система оценки качества представлена и проиллюстрирована результатами применения ARDS Объединенной Республики Танзания. Проверка качества данных с помощью выборочного обследования также обсуждается.

РИСУНОК 3. ДВА АСПЕКТА ОЦЕНКИ КАЧЕСТВА АДМИНИСТРАТИВНЫХ ДАННЫХ



ГЛАВА 3

ПОВЫШЕНИЕ КАЧЕСТВА АДМИНИСТРАТИВНЫХ ДАННЫХ ПО СЕЛЬСКОМУ ХОЗЯЙСТВУ

Поскольку административные данные собирают для достижения конкретных целей административного органа, а не для получения оценок характеристик конкретной целевой совокупности, они не обязательно напрямую применимы для достижения целей статистических организаций. Поэтому строгий процесс контроля качества, совершенствование подходов к сбору, вводу и обработке данных, а также решение вопросов, связанных с людскими и финансовыми ресурсами, являются ключевыми аспектами улучшения административных данных перед их использованием в системах сельскохозяйственной статистики.

3.1. МЕТОДОЛОГИЧЕСКИЙ ИНСТРУМЕНТАРИЙ ДЛЯ РАССМОТРЕНИЯ ВОПРОСОВ КАЧЕСТВА

Ивиг и др. (Iwig *et al* (2013) приводят руководящие принципы для координации действий с административными учреждениями для лучшего понимания того, как величины, полученные в результате административных процессов, соотносятся с характеристиками, представляющими интерес для статистических организаций.

В этом разделе основное внимание уделяется методологическим вопросам, а не вопросам структуры, действий и эффективности. Рассмотрены методологические вопросы, связанные с контролем качества, несколькими источниками данных, сбором данных, а также хранением и распространением данных; при этом отмечается необходимость обеспечения гармонизации и согласованности между различными наборами данных для обеспечения сопоставимости.

3.1.1. Контроль качества

Недостаток контроля над процессом сбора административных данных может привести к несовместимым форматам данных. НОСС может не иметь никакого контроля над сбором данных, и разные процессы сбора данных могут использовать разные форматы. Кроме того, административные органы могут со временем изменять протоколы сбора данных или использовать бумажные вопросники, что может привести к ошибкам обработки при оцифровке информации.

НОСС, отвечающий за сельскохозяйственную статистику, может оказать полезную поддержку производителям административных сельскохозяйственных данных, чтобы улучшить контроль, как над сбором данных, так и над их вводом. Независимые профессиональные статистики могут также внести свой вклад, предоставив диагностику проблем качества, касающихся процесса сбора данных, и технические рекомендации по улучшению. Однако наилучшим решением будет прием на работу статистиков этими административными учреждениями.

Системы административной отчетности для сельского хозяйства (ARSA)

В ARSA протокол контроля качества должен быть разработан таким образом, чтобы направлять мониторинг сбора данных, разработки данных и процессов передачи данных через систему от низового уровня до национального уровня. Например, протокол должен установить, когда данные должны быть собраны; когда супервайзеры должны проверить сборщиков данных и какова их ответственность; когда супервайзеры должны представлять отчеты в районные отделения; когда районные чиновники должны представлять отчеты в региональные отделения; и т.д. Должны быть установлены и соблюдены графики проведения консультаций или встреч для обсуждения вопросов, возникающих в ходе процессов передачи данных. Это обеспечит преемственность и устойчивость.

Усилия по более тщательному мониторингу процесса сбора административных данных также могут привести к улучшению качества данных. Galmes (2013) дает рекомендации по улучшению качества данных, полученных в результате процессов сбора данных ARSA, в том числе:

- использование стандартизированного формата для сбора информации;
- подготовка руководств, содержащих четкие определения деятельности;
- периодическое обучение сборщиков данных; а также
- строгий надзор.

Внедрение таких стандартов контроля качества при сборе данных также может повысить прозрачность процедур сбора данных. Это не только повышает качество данных; это также помогает пользователям (как статистическим аналитикам, так и потребителям статистических продуктов) в интерпретации данных и понимании того, как правильно их использовать.

Система традиционных административных данных

Процедуры контроля качества для оценки ошибок измерения и проблем охвата необходимы для защиты от смещения. Ошибки измерения связаны с процессом сбора данных, и проблема охвата может быть значительной, если целевая совокупность административных данных отличается от совокупности данных обследования.

Ошибки измерения в административных данных возникают по нескольким причинам. Понятийные различия часто существуют между величинами, собранными в ходе административных процессов, и величинами, представляющими интерес для статистического агентства. Например, административные процессы предполагают сбор информации о бенефициарах страхования от безработицы, однако определение безработицы, принятое в таких процессах, может отличаться от определения, используемого статистическими агентствами. Неверная отчетность также может быть вызвана различными мотивами в административных процессах. Например, фермеры могут занижать данные о площадях в программах субсидирования, чтобы избежать последствий непреднамеренного завышения (Carfagna and Carfagna, 2010). Ошибки измерения в определении переменных могут произойти, если заведения меняются; однако идентификатор, такой как адрес, не меняется. Изменения в характере административных процессов могут также привести к изменениям в собранных данных с течением времени, что затрудняет проведение последовательного лонгитюдного анализа.

Ошибки охвата возникают, когда совокупность, которая участвует в административном процессе, отличается от совокупности, представляющей интерес. Это может привести как к чрезмерному, так и к недостаточному охвату. Carfagna и Carfagna (2010) описывают исследования, проводимые для изучения проблем охвата в данных IACS. Они заключают, что если процедуры контроля качества указывают на существенные проблемы охвата, административные данные следует использовать только для поддержки данных выборочного

обследования, выбранных из совокупности для выборки, такой как AF, которая охватывает всю совокупность. Wallgren и Wallgren (2010) сравнивают проблемы охвата в реестрах предприятий и фермерских хозяйств и обнаруживают, что различия в свойствах охвата связаны с различными формами ошибок охвата.

Некоторые процедуры контроля качества для ошибок измерения и ошибок охвата могут быть, например, такими:

- База данных IACS поддерживается с целью управления программами помощи фермерам в ЕС. При ведении IACS, которая является административной базой данных для сельского хозяйства, проводится выборочный отбор деклараций на ежегодной основе, которые проверяются на наличие неправильностей, таких как ошибки действия или бездействия (Carfagna and Carfagna, 2010);
- ESSnet-ISAD (2008b) представляет собой тематическое исследование, иллюстрирующее использование деревьев решений для согласования определений переменных, связанных с пенсиями, из разных источников;
- Wallgren и Wallgren (2010) рекомендуют объединять несколько источников административных данных для улучшения охвата и проверки на наличие ошибок;
- Iwig *et al.* (2013) использует более упреждающий подход, предлагая руководящие принципы, предназначенные для того, чтобы помочь статистическим управлениям взаимодействовать с административными органами для согласования определений, унификации целей и улучшения своевременности обмена данными;
- Bakker (2012) разрабатывает подходы на основе моделей для квантификации смещения, вызванного ошибкой измерения в административных данных. Verka *et al.* (2012) исследуют эффективность теории Демпстера-Шефера для количественной оценки неопределенности в каждом элементе данных в каждом из нескольких регистров, используемых в переписи населения Австрии.

3.1.2. Улучшение сбора, хранения и распространения данных в рамках системы административной отчетности для сельского хозяйства (ARSA)

Использование передового опыта в сборе административных данных может сократить количество проблем качества, которые необходимо решить на этапе оценки. В системах традиционных административных данных (TAD) сбор данных обычно осуществляется с помощью хорошо отлаженных процедур и стандартизированных инструментов. Однако в системе административной отчетности для сельского хозяйства (ARSA) в процессах сбора, хранения и распространения данных используются различные методы и оборудование, и, возможно, потребуется улучшить их, чтобы повысить качество данных.

Совершенствование форм, вопросников и инструкций

Обеспечение того, чтобы используемый вопросник был как можно более коротким, может значительно улучшить качество собранных данных. Слишком длинный вопросник может накладывать излишнюю нагрузку на респондента, что приведет к неполучению ответов и низкому качеству данных. Этот широко известный принцип был подтвержден пилотным проектом, осуществленным в Объединенной Республике Танзания (GSARS, 2016a), поскольку стало очевидно, что даже вопросник, использованный в том проекте, мог бы быть пересмотрен или сокращен. Все инструкции должны быть включены в сборник инструкций, при этом должны быть отдельные руководства для счетчиков и супервайзеров. Это, в частности, сделает вопросник менее громоздким. В любом случае при использовании планшета инструкции могут быть запрограммированы внутри инструмента и легко открываться отдельно. Еще один способ улучшения качества сообщаемых данных состоит в том, чтобы информировать респондентов о значимости и предполагаемом использовании данных; это должно быть описано в инструкциях.

Использование новых технологий при сборе данных

Использование определенных технологий при сборе административных данных может способствовать повышению качества данных, особенно с точки зрения своевременности. Использование оборудования GPS в измерениях расстояний и площадей повышает точность при одновременном сокращении времени, необходимого для сбора данных, хотя это также может увеличить затраты на сбор информации по сравнению с декларациями фермеров (что влияет на устойчивость ADSAS).

Еще одна важная технология, которая все чаще используется для сбора данных, - это программное обеспечение для интервьюирования с помощью компьютера (CAPI) через мобильные телефоны или планшеты. Эта современная, но дешевая и легкодоступная технология может улучшить своевременность сбора данных. На местах ввод данных будет осуществляться сборщиками данных (например, работниками по распространению сельскохозяйственных знаний в деревне).

Если используются электронные устройства с CAPI, данные записываются непосредственно в базу данных, которая доступна на районном, региональном или провинциальном и национальном уровнях. Права доступа должны быть предоставлены таким образом, чтобы супервайзеры могли проверять данные, а затем направлять их районным специалистам, которые, в свою очередь, проверяют и очищают их для доступа на более высоких уровнях. Как отмечалось в ходе экспериментального проекта, проведенного в Кот-д'Ивуаре, счетчикам следует предоставить некоторое количество бумажных вопросников, которые будут использоваться в случае отказа устройства (GSARS, 2016a).

Программное обеспечение CAPI должно быть спроектировано таким образом, чтобы отчеты могли автоматически генерироваться после разработки данных на районном уровне. Эти отчеты могут использоваться в качестве основы для мониторинга системы, а также для пересылки данных на более высокие уровни и для целей районного планирования.

Во время пилотного проекта было также продемонстрировано, что предоставление фермерам «карты урожая» может улучшить оценку производства, особенно для смешанных, многолетних и постоянно собираемых культур. Во время пилотного проекта было также предложено запрограммировать карту урожая, используемую при оценке урожая, на мобильных телефонах, чтобы фермеры могли отчитываться регулярно и в электронном виде. Тем не менее, было рекомендовано ограничить использование карт урожая для непрерывно собираемых культур.

Улучшение передачи и представления данных

Если в районе или регионе имеется достаточный уровень подключения к Интернету, данные могут быть немедленно переданы на центральный сервер, как только они введены в приложение CAPI на планшете. Хотя данные могут передаваться при наличии подключения к Интернету, данные и отчеты должны рассматриваться как официально представленные только после того, как они были утверждены супервайзерами. Если нет подключения к Интернету, то данные временно сохраняются на планшете и отправляются на сервер, как только соединение восстанавливается.

Улучшение хранения и распространения данных

Это охватывает информационно-коммуникационные технологии (ИКТ) и носители, используемые при хранении и распространении данных: портативное оборудование, телефоны, традиционные средства связи (такие как радио, телевидение и факс); современные ИКТ (например, электронная почта, интернет и SMS); КПК и т.п.

Как обсуждалось в GSARS (2015a), метаданные имеют жизненно важное значение для информирования производителей и пользователей о качестве данных. Рекомендуется, чтобы метаданные были представлены на всех этапах. Входящие данные должны сопровождаться достаточным количеством метаданных, чтобы обеспечить их полное понимание и правильное распределение значений для соответствующих переменных. Метаданные являются основой управления показателем интерпретируемости. Примером является Интегрированная база метаданных (IMDB) - единый источник метаданных Статистического управления Канады, описывающий обследования и программы. Качество информации в IMDB должно регулярно контролироваться для обеспечения полноты и точности. Было подчеркнуто, что статистическим агентствам важно публиковать хорошие метаданные, поскольку при этом они демонстрируют открытость и прозрачность и, таким образом, способствуют повышению доверия среди пользователей данных (Dion, 2007).

При внедрении новых технологий хранения или распространения данных учреждениям рекомендуется учитывать связанные с этим преимущества и риски. Новая технология может принести выгоду в определенных измерениях, в то же время создав затраты в других областях. Вопросы, которые необходимо учитывать в отношении надежности, доступности, своевременности и устойчивости, заключаются в следующем.

Надежность: может ли технология повысить точность распространяемой информации? Использование некоторых технологий, таких как SMS или электронная почта, уменьшает количество ошибок распространения.

Доступность: Некоторые неграмотные пользователи не смогут прочитать SMS и электронные письма.

Своевременность. Некоторые технологии могут обеспечить более быструю передачу административных данных и информации (например, информация о ценах может быть быстрее отправлена по SMS по сравнению с обновлением веб-сайта).

Устойчивость. Этот критерий касается затрат, связанных с использованием технологий для хранения или распространения информации. Некоторые технологии могут быть быстрыми, но дорогими, например планшеты iPad, что связано с высокими постоянными затратами. Кроме того, некоторые технологии могут оказаться невозможными, если возникают проблемы с электричеством и безопасностью.

3.2. ЛЮДСКИЕ И ФИНАНСОВЫЕ РЕСУРСЫ

3.2.1. Людские ресурсы

Системы административной отчетности для сельского хозяйства (ARSA)

Необходимо провести оценку потребностей в людских ресурсах, чтобы определить, имеется ли недостаток кадров на каком-либо уровне рутинной системы отчетности, от национального до местного уровня управления. Объем финансирования, необходимый для найма и обучения персонала, должен быть определен и включен в национальный бюджет и бюджет административного органа (полугосударственных и местных органов власти). Подбор необходимого персонала должен осуществляться на основе выявленных пробелов в кадрах.

Иерархический протокол обучения должен быть разработан для использования на всех уровнях. На национальном уровне служащие должны быть обучены. Они должны обучать региональных служащих, которые, в свою очередь, обучают районных специалистов по распространению знаний. Руководства по обучению должны содержать подробную информацию о том, как использовать инструменты сбора данных на низовом уровне, а также формат сбора данных на

районном, региональном и национальном уровнях. Существует также необходимость в четких пособиях по процессам сбора, анализа и распространения данных. Также должен проводиться систематический контроль со стороны квалифицированного персонала, в том числе, если это возможно, сотрудников НСУ.

ПРИМЕР ХОРОШЕЙ ОРГАНИЗАЦИИ ОБУЧЕНИЯ - ARDS ОБЪЕДИНЕННОЙ РЕСПУБЛИКИ ТАНЗАНИЯ

- Для районных специалисты предоставляется учебное пособие по консолидации данных, анализу и обратной связи в ARDS, которое содержит рекомендации по обработке и анализу данных на уровне районов.
- Региональные специалисты и районные специалисты проходят обучение по общим формам отчетности.
- Районные специалисты, в свою очередь, обучают сотрудников по распространению сельскохозяйственных знаний на уровне деревень и участков формату сбора данных на этом уровне.
- Тренинг по Excel и базе данных мониторинга местных органов власти 2 (LGMDB2) для управления данными:
 - Региональные специалисты и ИТ-специалисты из нескольких соседних регионов собираются и проходят обучение в одном месте по Excel и LGMDB2;
 - Региональные специалисты затем обучают районных сотрудников Excel и LGMDB2.

Система традиционных административных данных

Для этой категории административных данных может осуществляться передачу знаний. Профессиональные статистики или независимые статистические эксперты в НСУ могут изучить имеющиеся инструменты и процессы для сбора информации и провести обучение или предоставить рекомендации по их улучшению. Обмен персоналом является еще одной практикой, которая может быть реализована. Например, Статистическое управление Канады и Налоговая служба Канады осуществляли такой обмен кадрами в целях улучшения качества получаемых налоговых данных.

3.2.2. Финансовые ресурсы

Учреждения, отвечающие за производство административных сельскохозяйственных данных, нуждаются в особой финансовой поддержке в части затрат на сбор данных. В развивающихся странах это обычно не так просто из-за бюджетных ограничений. Тем не менее, это вопрос, который может существенно повлиять на качество данных и является особенно важным для ARSA. Лучшая координация между учреждениями, собирающими сельскохозяйственную статистику, могла бы минимизировать затраты на процессы сбора, разработки и управления данными. Это может быть дополнительно улучшено за счет более тесного сотрудничества с пользователями данных. Для конкретного случая координации систем административной отчетности могут быть полезны следующие рекомендации:

- Пропаганда передового опыта - лоббирование правительств в некоторых странах Азиатско-Тихоокеанского региона доказало свою успешность в обеспечении доступа к национальным бюджетам для целей системы сбора рутинных данных. Поэтому правительства должны быть осведомлены о важности и преимуществах наличия функциональной рутинной системы сбора данных. Даже местные органы власти (административные единицы) следует убедить выделять средства для обеспечения устойчивости обычной системы сбора данных;
- Объединение ресурсов (людских и финансовых) и гармонизация административных данных различных учреждений может сократить расходы в странах. Однако при этом роли и обязанности каждого учреждения (особенно министерства, отвечающего за сельское

хозяйство и НСУ) должны быть четко изложены в меморандуме о взаимопонимании (МОВ) или в другом правовом документе.

3.3. СТАНДАРТИЗАЦИЯ ПОНЯТИЙ В УЧРЕЖДЕНИЯХ, ЗАНИМАЮЩИХСЯ АДМИНИСТРАТИВНЫМИ ДАННЫМИ В СЕЛЬСКОМ ХОЗЯЙСТВЕ

Несколько применений административных данных для производства официальной статистики подразумевает интеграцию источника административных данных либо с другими источниками административных данных, либо с обследованиями или переписями. Проблемы в этом процессе интеграции возникают, когда разные источники данных используют разные определения или системы кодирования. Например, простые различия в метках, используемых для идентификации единиц в микроданных, могут препятствовать связыванию разнородных источников данных. Отсутствие стандартизации означает, что существуют несоответствия между формами сбора данных. Когда процессы децентрализованы и не стандартизированы, разные методы сбора данных могут использовать разные форматы, что может затруднить интеграцию источников данных.

Процесс стандартизации понятий между учреждениями ADSAS должен учитывать различия между (i) определениями единиц (таких как фермерские хозяйства), (ii) определениями переменных (например, временных сотрудников) и (iii) системами кодирования. В этом разделе будет подчеркнута необходимость этого процесса. Эти согласования могут быть выполнены посредством организации специальных консультаций или совещаний технических групп.

3.3.1. Гармонизация определений единиц

В данном контексте единица – это самая мелкая отчитываемая единица в файле микроданных. Определения единиц в административных файлах, как правило, определяются функцией административного органа, о котором идет речь. Следовательно, определение единицы административным органом может, например, отличаться от определения, одобренного НСУ. Многочисленные индивидуальные фермы могут составлять одну операцию; поэтому единица в административной базе данных может отличаться от статистической единицы.

Бенедетти и др (Benedetti *et al.* (2010) считают, что проблема, связанная с единицами, настолько существенна, что является главной заботой для статистиков. Широкий выбор возможных единиц (семья, фермерское хозяйство, домашнее хозяйство, земельный участок, пункт и т.д.) и зависимость от наличия качественной совокупности единиц являются особо важными вопросами. Различия между определениями единиц могут ограничивать полезность административного источника, особенно для целей, связанных с увязкой микроданных.

Поэтому необходимо гармонизировать определения единиц, принятых учреждениями ADSAS и НСУ.

3.3.2. Гармонизация идентификаторов единиц

Важно, чтобы в административных источниках данных применялся одинаковый подход к идентификации единиц. Переписи и обследования, проводимые в стране, должны в максимально возможной степени использовать одни и те же идентификаторы. Это облегчит увязку баз данных для различных видов использования, описанных в главе 5.

3.3.3. Гармонизация определений переменных

Различия в определениях родственных понятий могут привести к значительным несоответствиям между административными и статистическими источниками. Например, определение дохода для целей налогообложения может отличаться от того, которое требуется лицам, определяющим политику, и которые хотят проанализировать данные в контексте обследования доходов. Такие различия могут вызывать систематические отклонения между величинами, полученными из административного источника, и соответствующими величинами (или их оценками), полученными на основе обследований (ООН, 2011; Wallgren and Wallgren, 2010; Carfagna and Carfagna, 2010; Brackstone, 1987). Определения и содержание административных записей иногда изменяются без предварительного уведомления пользователей и без предоставления переходного периода, в который новые и старые определения используются одновременно. В отсутствие периода, в течение которого данные собираются в соответствии с обоими определениями, невозможно выявить реальное изменение, так как на оценки мог повлиять пересмотр определений. Влияние изменений в определениях более заметно при непрерывном обновлении файлов, как это может быть в случае с регистром (UNESC, 2007).

В случае изменений в определениях определения переменных должны быть согласованы, и должна осуществляться систематическая связь с НСУ.

3.3.4. Гармонизация системы кодирования

Проблемы, возникающие из-за использования различных систем кодирования, тесно связаны с проблемами, связанными с различиями в определениях переменных и единиц. Например, НСУ может потребоваться более детальная система кодирования, чем та, которая требуется административному органу. При объединении источников с различными системами кодирования возникают неоднозначные ситуации, когда один код в административном источнике увязывается с несколькими кодами в системе, используемой НСУ.

РЕЗЮМЕ

В этой главе рассматриваются различные подходы, которые могут быть приняты для повышения качества административных данных. Описаны методологические инструменты для повышения качества данных, которые охватывают сбор данных и контроль качества, а также хранение и распространение данных. Также обсуждаются вопросы людских и финансовых ресурсов и стандартизации понятий среди административных учреждений, занимающихся сельскохозяйственными данными. Рекомендуется вводить проверки контроля качества на всех этапах и стандартизировать концепции путем согласования определений единиц и переменных, а также систем кодирования, принятых в различных административных источниках данных. Что касается ARSA, то инструменты сбора данных должны быть улучшены, и должны быть предоставлены адекватные финансовые и людские ресурсы.

ГЛАВА 4

УЛУЧШЕНИЕ ДОСТУПНОСТИ АДМИНИСТРАТИВНЫХ ДАННЫХ

Доступ к административным данным может быть ограничен правовыми и политическими ограничениями, которые могут иметь место не без причин - например, для защиты конфиденциальности отдельных лиц (Brackstone, 1987).

Для облегчения доступа к административным данным в статистических целях необходимо развитие большого количества систем. Эти системы, как правило, имеют несколько измерений: юридические, политические, организационные и технические. Необходимо достичь соглашения во всех этих областях, прежде чем преимущества использования административных данных могут быть полностью достигнуты.

4.1. ПРАВОВАЯ СИСТЕМА

ООН (2011) обсуждает значение законов и политики для обеспечения доступа статистических управлений к необходимым административным данным. Во многих случаях существует законодательство, прямо предусматривающее доступ к административным данным. Например, Законы о статистике Ирландии и Норвегии устанавливают разрешение для НСУ соответствующей страны на доступ к административным данным. В Законе о статистике Ирландии 1993 года говорится: «с целью оказания помощи [Статистическому] ведомству в выполнении его функций в соответствии с настоящим Законом Генеральный директор может путем направления уведомления запрашивать любой государственный орган (а) разрешить сотрудникам статистических служб в любое разумное время иметь доступ к проверке, а также копированию или получению выдержек из любых записей, находящихся в его ведении, и (б) предоставлять ведомству, если того требует такой сотрудник, копии или выдержки из любой такой записи и [...] государственный орган обязан выполнить любой такой запрос бесплатно» (UN, 2011). Поскольку возможности для принятия законодательства ограничены, а эффективное воздействие на правовую систему требует значительных усилий, статистическим управлениям рекомендуется предлагать законодательство с учетом долгосрочной стратегии (ООН, 2011 г.). В примере, взятом в Статистическом управлении Канады, Брэкстоун (Brackstone, 1987) отмечает, что налоговая реформа правительства - подходящее время для статистического управления взаимодействовать с политиками и стремиться формировать данные, собираемые с помощью правовых структур, таким образом, чтобы удовлетворять потребности статистического управления. Международные стандарты также помогают с точки зрения предоставления руководящих указаний, и поэтому на них следует ссылаться, когда это возможно, в ходе обсуждений с административными органами.

Правовые рамки обычно формируются на национальном уровне и характерны для национальных источников и обстоятельств. В некоторых случаях, однако, может также существовать соответствующее законодательство на международном уровне. В этих случаях могут быть две или более альтернативных правовых возможностей для использования административных данных.

Большинство НСУ имеют правовую основу, определяющую их роли и обязанности, как правило, в форме закона о статистике. В некоторых странах эти правовые документы были пересмотрены в последние годы и в настоящее время включают конкретные положения, обеспечивающие доступ к административным данным. Страны, которые не внесли такие поправки, должны сделать это, поскольку это является необходимым шагом.

Национальные исторические, политические и институциональные факторы оказывают сильное влияние на эти правовые рамки. В результате могут возникнуть национальные различия, которые приведут к созданию правовых систем, которые не особенно согласованы или даже не согласованы между странами. Международная сопоставимость статистических данных, которые были получены полностью или частично из административных источников, может быть улучшена с помощью международной правовой базы, регулирующей доступ к административным данным.

Помимо обеспечения доступа к данным из административных источников, правовые системы должны также устанавливать ограничения для такого доступа и возможного использования административных данных. Зачастую существуют ограничения, согласно которым данные могут использоваться только для конкретных статистических целей, или же следует сохранять конфиденциальность индивидуальных данных. Также могут быть определенные ограничения на использование данных.

4.2. НОРМАТИВНАЯ БАЗА

Когда внесение изменений в правовые системы нецелесообразно, могут быть сформулированы меры, облегчающие доступ к административным данным или их изменение. Меры легче изменить, чем законы, и они имеют тенденцию развиваться более динамично с течением времени (ООН, 2011). Одним из примеров нормативной базы, касающейся административных данных, является Принцип 5 «*Основопологающих принципов официальной статистики*» ООН, который подчеркивает экономическую эффективность административных данных и способствует использованию таких источников в интересах эффективного использования имеющейся информации (ООН, 2011). Этот принцип из документа, утвержденного Генеральной Ассамблеей ООН, может использоваться НСУ для пропаганды расширения доступа к административным данным в статистических целях.

Во многих странах существуют общие правила обмена данными внутри государственных органов, которые будут влиять на право доступа к административным данным в статистических целях. Нормативная база также включает добровольные кодексы норм, наиболее важным из которых для статистических целей являются *Основопологающие принципы официальной статистики ООН*.

Кодексы норм также должны публиковаться на национальном уровне, чтобы заверить общественность в том, что данные будут использоваться только для конкретных целей. Чтобы эти кодексы норм имели реальную ценность, важно, чтобы они были доступны для широкой публики.

После того, как будут созданы правовые и нормативные базы, позволяющие использовать административные данные, необходимо рассмотреть организационные механизмы для облегчения потоков данных. Как правило, это принимает форму письменного соглашения или меморандума.

4.3. ВОПРОСЫ, СВЯЗАННЫЕ И КОНФИДЕНЦИАЛЬНОСТЬЮ И ВОСПРИЯТИЕМ ОБЩЕСТВЕННОСТИ

Вопрос конфиденциальности является сложным, особенно в отношении административных данных. Рассмотрим два различных сценария: один, в котором некоторая информация в административной базе данных не защищена законами, обеспечивающими секретность, и другой, в котором административная информация является конфиденциальной. В первой ситуации проблема конфиденциальности не является препятствием для обмена данными с политической или юридической точки зрения. Во втором случае соглашение между статистическим управлением (часто НСУ) и административным органом необходимо для обеспечения доступа к административной информации. Это разрешение может принимать форму меморандума о взаимопонимании, переопределения статистической системы или правительственного законодательного акта или мер, позволяющих статистическому управлению получать доступ к административным данным.

Во втором сценарии, описанном выше, частные лица и предприятия могут предоставлять информацию административному органу, понимая, что сообщаемая информация останется конфиденциальной. Следовательно, использование административных данных в статистических целях может быть встречено со скептицизмом со стороны общественности (Brackstone, 1987).

Чтобы удовлетворить этот общественный запрос на конфиденциальность, статистическому управлению рекомендуется принять адекватные меры для обеспечения конфиденциальности административных данных. Например, в Статистическом управлении Канады административные данные о налогах размещаются в строго ограниченной и безопасной области. Необходимость учитывать озабоченность общественности в отношении конфиденциальности и сохранности персональных данных может увеличить расходы, связанные с административными данными, и ограничить их доступность.

Страны, сталкивающиеся с проблемами доступа к некоторым административным данным из-за ограничений конфиденциальности, могут тщательно пересмотреть свое законодательство о статистике и сравнить свои положения о конфиденциальности с положениями, установленными в нормативном законодательстве производителей административных данных. Если положения Закона о статистике идентичны или более строгие с точки зрения конфиденциальности, НСУ может говорить о том, что получение доступа к административным данным не представляет риска нарушения конфиденциальности.

4.4. СОГЛАШЕНИЯ МЕЖДУ УЧРЕЖДЕНИЯМИ

При имеющихся правовых и нормативных базах, необходимых для разрешения использования административных данных, часто необходимы письменные соглашения для детализации и облегчения передачи знаний и данных (ООН, 2011). Эти письменные соглашения часто имеют форму меморандума о взаимопонимании, который определяет цели статистического управления в использовании административных данных и информацию, необходимую для достижения этих целей (Prell *et al.*, 2009).

4.4.1. Обзор и преимущества меморандума о взаимопонимании (MoV)

Брэкстоун (Brackstone, 1987) обращает внимание на успех Статистического управления Канады в создании "двусторонних комитетов" при участии статистического управления и административных органов в развитии необходимой организационной и технической инфраструктуры.

Прелл и др. (Prell *et al.*, 2009) проанализировали семь практических примеров, включающих письменные соглашения, которые устанавливают или развивают отношения с административными органами. Они определяют четыре отличительные характеристики успешного МоВ:

- i. **видение и поддержка со стороны руководства агентства:** межведомственные проекты по обмену данными могут потребовать значительного участия со стороны руководства агентства. Однако статистическое использование данных агентства может рассматриваться руководством как второстепенная задача;
- ii. **узкие, но гибкие цели:** для МоВ может быть важно узко определить цели, иногда вплоть до того, значениями каких конкретных полей в базе данных будут обмениваться агентства и как эти поля будут использоваться. Узкие цели также полезны, потому что они способствуют межведомственному обсуждению вопросов управления данными. Хотя полезно, чтобы цели были достаточно узкими для плодотворного обсуждения в межведомственных дискуссиях, цели также должны быть гибкими, то есть изменяться в результате этих обсуждений;
- iii. **инфраструктура:** этот элемент успеха состоит из двух компонентов: кадровое обеспечение и меры и процедуры. Межведомственные проекты выгадывают от участия людей, ориентированных на результаты, поддерживающих цели проекта, имеющих опыт работы с данными и способных работать совместно с сотрудниками своего агентства и агентства-партнера. Количество рабочего времени сотрудников часто было проблемой для семи проектов обмена данными, рассмотренных в работе Prell *et al.* (2009). Вторым компонентом инфраструктуры является важность наличия соответствующих мер и процедур для поддержки деятельности по обмену данными;
- iv. **взаимный интерес:** чтобы достичь успеха, механизмы обмена данными должны приносить пользу каждому партнеру проекта.

Рассмотренные случаи указывают на то, что эти «элементы успеха» дают возможность агентствам решить многие проблемы, которые возникают в процессе подготовки письменного меморандума о взаимопонимании.

Ивиг и др. (Iwig *et al.*, 2013) предлагает схему взаимодействия между статистическим управлением и административным органом при формировании отношений по обмену данными. Их работа «Инструмент оценки качества данных для административных данных (*Data Quality Assessment Tool for Administrative Data*)» построен вокруг таких аспектов качества, как актуальность, доступность, согласованность, интерпретируемость, точность и институциональная среда. Для каждого измерения качества в работе Iwig *et al.* (2013) рекомендуется несколько вопросов, которые статистическое управление должно задать административному органу. Например, в интересах обеспечения согласованности понятий, классификаций и методов сбора данных во времени и по географическим областям следует сделать следующий запрос: «Пожалуйста, опишите все системы классификации, используемые для категоризации или классификации данных».

Организационные соглашения также могут преодолеть ограничения, связанные с сохранением конфиденциальности. Если и статистические данные, и административные данные считаются конфиденциальными, меморандум о взаимопонимании, как обсуждалось в Prell *et al.* (2009) может определить правовой механизм для передачи данных. В некоторых случаях расширение определения национальной статистической системы может способствовать более либеральному распространению административных данных среди государственных учреждений, включая статистические управления. Как обсуждали Wallgren и Wallgren (2007), Статистическое

управление Швеции регулярно получает административные данные от агентств, отвечающих за государственные программы и регулирование.

Административные данные должны собираться с уведомлением о том, для каких целей будет использоваться эта информация, с тем, чтобы были ясны обстоятельства, когда административные записи следует рассматривать как конфиденциальную информацию и обращаться с ней соответственно.

Обмен данными между агентствами касается тех методов, с помощью которых ведомства могут получать доступ к данным друг друга о лицах, иногда сразу, но почти всегда своевременно. Обмен данными предлагает ряд преимуществ. Если разные агентства собирают одинаковые данные об одном и том же лице, процесс сбора дублируется как для ведомств, так и для человека. Таким образом, обмен данными может повысить эффективность за счет снижения бремени бумажной работы для государственных органов и для отдельных лиц, поскольку базовую информацию о клиентах необходимо получить только один раз. Также возможно улучшить процент ответов.

Хотя обмен данными имеет много преимуществ, он поднимает вопросы, касающиеся защиты личных данных и конфиденциальности: кто должен иметь доступ к этим данным; как конфиденциальность и права на защиту личных данных могут быть обеспечены при достижении выгод от увязки программных данных; и т.д. Все эти вопросы должны быть учтены при разработке МОВ.

4.4.2. Основные моменты для включения в МоВ

a. Правовая основа.

Следует сделать ссылку на законодательство, разрешающее доступ к административному источнику для статистических целей, и на любое законодательство, которое налагает ограничения на такой доступ.

b. Имена лиц, передающих и получающих данные.

Имена и контактные данные ключевых людей, участвующих в предоставлении данных в административных и статистических организациях, должны быть записаны.

c. Подробное описание охвата данных.

Это будет включать информацию, идентифицирующую набор данных и содержащиеся в нем переменные.

d. Частота предоставления данных.

Это будет указывать, когда и как часто административный орган будет предоставлять необходимые данные.

e. Стандарты качества.

Здесь устанавливаются параметры качества поставляемых данных. Примером может быть указание максимально допустимой доли отсутствующих или ошибочных переменных, чтобы гарантировать, что полученные данные соответствуют цели. Приоритеты, присвоенные различным переменным, и, следовательно, усилия, предпринимаемые для обеспечения качества, часто различаются в административных и статистических организациях; поэтому согласование общих стандартов имеет первостепенное значение.

f. Правила обеспечения конфиденциальности.

Важно четко указать, как могут быть использованы данные, какие существуют правила и процедуры для предотвращения их раскрытия, а также обстоятельства, при которых данные могут передаваться клиентам статистической организации.

g. Технические стандарты.

Это измерение включает в себя следующие аспекты:

- Предоставление метаданных;
- Важно, чтобы потоки данных сопровождалась соответствующими метаданными, которые могут включать даты, описания используемых кодов, информацию об используемых единицах и пр.;
- Плата за поставку данных;
- Передача данных между государственными ведомствами или агентствами, как правило, осуществляется бесплатно, хотя в некоторых случаях от статистической организации может потребоваться участие в расходах на извлечение и передачу данных. За данные от организаций частного сектора может взиматься плата по рыночным ставкам, хотя может быть возможным договориться о скидках, особенно если среди органов государственного управления есть несколько пользователей источника данных частного сектора. В некоторых случаях может быть возможно предложить проведение статистического анализа или экспертизы в качестве формы оплаты за полученные данные;
- Срок соглашения;
- Обычно соглашения заключаются на фиксированный период, но должны включать возможность возобновления или продления при необходимости;
- Непредвиденные обстоятельства в связи с изменением условий;
- Для статистической организации важно заранее получать предупреждение об изменениях, касающихся административного источника. В соглашении должно быть указано, что о любых предлагаемых изменениях следует сообщать статистической организации как можно скорее, чтобы свести к минимуму влияние изменений на статистические результаты;
- Процедура для разрешения споров;
- В соглашении должен быть указан метод, который следует использовать при разрешении любых споров, которые могут возникнуть между статистическими и административными организациями; они могут предусматривать участие руководителей или, возможно, даже соответствующих министров.

h. Техническая база.

Техническая база - это механизмы, с помощью которых передаются данные, а также любые соответствующие стандарты данных или метаданных. Принятый механизм передачи данных должен учитывать технические возможности, доступные как отправляющей, так и принимающей организации.

Пример Меморандума о взаимопонимании приведен в Приложении 4 к данному Руководству.

РЕЗЮМЕ

В этой главе обсуждаются способы улучшения правовых и нормативных баз для облегчения доступа к административным данным в статистических целях. Даются рекомендации по поводу решения вопросов, связанных с конфиденциальностью и общественным восприятием, а также по поводу заключения соответствующего соглашения между учреждениями, а именно меморандума о взаимопонимании. Рекомендуется:

- обеспечить наличие адекватных правовых и нормативных баз;
- решать вопросы конфиденциальности; а также
- заключить соглашение между организациями, которые бы обеспечивали учет ключевых особенностей.

ГЛАВА 5

ИСПОЛЬЗОВАНИЕ АДМИНИСТРАТИВНЫХ ДАННЫХ

Административные данные часто собираются с высокой частотой и детальной географической разбивкой. Поскольку административные данные по определению собираются для нестатистических целей, статистическое агентство несет относительно низкие затраты на сбор данных при их использовании. Это позволяет использовать административные данные для нескольких целей.

Обзор, проведенный в GSARS (2015a, 2015b), показал, что использование административных данных можно разбить на две большие категории: косвенное использование и прямое использование. При косвенном использовании административные данные применяются при разработке или совершенствовании статистического продукта, который также использует данные обследований или переписей. Прямое использование относится к ситуациям, в которых административная информация используется в качестве конечного статистического продукта для независимых целей, таких как государственное планирование. Статистические управления развитых стран используют административные данные как прямо, так косвенно. Развивающиеся страны с большей вероятностью будут напрямую использовать такие данные, особенно когда финансовые ограничения мотивируют применение административных данных для замены данных обследований или переписей.

Эффективное использование административных данных требует понимания множества аспектов качества данных и «пригодности для использования» при производстве официальной статистики. Эти методологические вопросы должны решаться на практике при использовании административных данных в качестве прямого источника информации или косвенно при улучшении итогового статистического продукта.

В этой главе приводятся примеры прямого и косвенного использования административных данных и предоставляются методологические инструменты для решения проблем качества, отмеченных в предыдущих главах. Обычной операцией, которая выполняется до эффективного использования административных данных, является увязка записей, которая служит для интеграции данных из различных административных источников или административных данных и данных переписи или обследования. Поэтому глава начинается с описания основных инструментов, используемых для увязки записей. В следующих разделах исследуются использование административных данных при разработке статистического продукта и концепция прямого использования.

5.1. МЕТОДОЛОГИЧЕСКИЕ И ТЕХНИЧЕСКИЕ ИНСТРУМЕНТЫ ДЛЯ УВЯЗКИ ЗАПИСЕЙ И ИНТЕГРАЦИИ ДАННЫХ

Операция по интеграции данных представляет вычислительные и методологические проблемы. Статистическое программное обеспечение может уменьшить вычислительную нагрузку при

объединении больших наборов данных. Статистические методы, такие как профилирование и вероятностная увязка записей, могут справиться с отсутствием стандартизации в определениях единиц или идентификации переменных. В зависимости от целей и имеющихся данных, интеграция нескольких источников может осуществляться на уровне отдельной единицы или для агрегированной группы единиц (то есть региона или страны). Ниже рассмотрены примеры методов интеграции на уровне единиц и на уровне агрегатов. Более подробная информация доступна в GSARS (2015b и 2017b).

5.1.1. Профилирование

Одной из проблем, связанных с интеграцией нескольких источников информации, является то, что разные источники данных могут иметь разные определения единиц. В материале ООН (2011) объясняется, что «... преобразование административных единиц в статистические единицы может быть довольно сложным в теоретическом плане и часто включает некоторую форму моделирования». Термин «профилирование» используется в бизнес-обследованиях для описания этого процесса, хотя это понятие применяется и в других контекстах (ООН, 2011).

Профилирование может проводиться в ручном или автоматическом режиме (ООН, 2011). Стандартные правила, основанные на атрибутах или характере связей между единицами, могут помочь преодолеть различия между административными и статистическими единицами. Например, в статистике домохозяйства могут быть сформированы на основе отношений между людьми, живущими в одном здании; действительно, этот подход является компонентом метода переписи населения на основе регистров, используемого в скандинавских странах. Даже при профилировании силами офисных работников разукрупнение единиц может потребовать субъективных определений, и единого правильного решения может не существовать. В автоматизированных процессах, которые дешевле и быстрее, чем профилирование силами офисных работников, стандартные правила, касающиеся характера связей, применяются единообразно.

Альтернатива профилированию на основе правил включает построение статистических моделей. Отношения между административными и статистическими единицами могут быть установлены для подгруппы совокупности, например, посредством обследования; а параметры моделей, описывающих взаимосвязи, затем оценены, и их значения распространены на всю совокупность. Примером может служить случай, когда административная единица – это «рабочее место», а статистическая единица – «лицо» (ООН, 2011). В оценке, основанной на опросе, на каждого человека приходится в среднем 1,15 рабочих мест; эта оценка может использоваться в качестве глобального корректирующего фактора для определения оценки количества занятых на основе количества рабочих мест. Изменчивость оценки этого фактора, основанной на результатах обследования, должна быть включена в последующий анализ занятости.

5.1.2. Детерминистическая увязка записей

Одним из механизмов улучшения охвата и уменьшения ошибки измерения является интеграция нескольких административных источников для формирования систем регистров. Этот процесс интеграции требует увязки единиц из разных файлов. Когда в разных файлах используются уникальные идентификационные номера, детерминистическая увязка может быть выполнена посредством простого слияния. Поскольку базы административных данных могут содержать различные типы единиц, соотношения между единицами не обязательно являются «один-к-одному», и необходимы процедуры для построения связей «многие-к-одному» или «один-ко-многим» (Wallgren and Wallgren, 2010).

5.1.3. Вероятностная увязка записей

Многие применения административных данных для производства официальной статистики предполагают наличия многих источников информации – много файлов административных данных или файлов административных данных и данных обследований. Для многих видов использования административных данных желательно проводить увязку записей из, как минимум, двух файлов на уровне отдельных единиц совокупности. Рассмотрим, например использование административных данных для проверки ошибок в данных обследования. Хотя сравнения безусловных распределений файла административных данных с соответствующими безусловными распределениями из обследования или переписи может быть информативным, сравнение других источников данных на уровне единиц открывает больше возможностей. Операция по увязке на уровне единиц, например, позволяет оценить записи с относительно большими различиями величин, зарегистрированных в двух разных источниках.

Операция по слиянию файлов на уровне отдельных записей представляет большие проблемы. Идентифицирующие переменные могут различаться в разных наборах данных. Даже если существует уникальный идентификатор, идентифицирующая переменная может отсутствовать или быть некорректно приведена для некоторых единиц. Дублирующие записи могут существовать в одном или более файлах. Для больших наборов данных могут требоваться значительные вычислительные усилия. Вероятностная увязка записей – это статистическая процедура для определения вероятности того, что два набора идентифицирующих переменных представляют одну и ту же единицу совокупности.

Феллеги и Сантер (Fellegi and Santer, 1969) разработали одну из самых широко используемых процедур для вероятностной увязки записей. В предложенной ими модели рассматриваемый латентный статус совпадения представлен латентной бинарной переменной, δ , то есть 1, если заданная пара совпадает, и 0 - в противном случае. Наблюдения – это векторы сопоставляемых переменных, $\gamma = (\gamma_1, \dots, \gamma_K)$, где $\gamma_j = 1$, если x_j^A из набора **A** равна переменной x_j^B из набора **B** и равна 0 в противном случае. Распределение, используемое для вывода, – это условное распределение δ при заданных наблюдаемых векторах сравниваемых переменных.

Было разработано много расширений для вышеуказанной процедуры. Например, Ларсен и Рубин (Larsen and Rubin, 2001) разрабатывают процедуру, которая включает в себя чередование умозаключений на основе условных распределений и ручного просмотра. Включение офисного просмотра снижает степень неопределенности. Когда наборы данных увязаны при помощи вероятностных моделей увязки или при неполной увязке, процедуры оценки должны учитывать ошибку увязки. Можно использовать такие методы, как были предложены Kim and Chambers (2012), которые распространяют методы регрессионного анализа Chambers (2009) на приложения с более чем двумя увязанными наборами данных. Verka *et al.* (2012) изучают эффективность теории Демпстера-Шафера (Dempster-Shafer) для количественной оценки неопределенности в каждом элементе данных в каждом из нескольких регистров, используемых для переписи на основе регистров в Австрии.

i. Шаги по выполнению очистки данных и увязки записей

Процесс подготовки набора данных для увязки записей и выполнения алгоритма увязки включает несколько взаимосвязанных компонентов. Один из них состоит в выборе переменных, используемых для сопоставления. Затем, при имеющемся наборе совпадающих переменных, часто бывает нужна операция по приведению представления в разных наборах данных к стандартной форме. Поскольку сравнение всех пар записей в двух файлах часто неприемлемо с точки зрения объема расчетов, необходимы методы для сокращения размерности пространства сравнения. При сравнении двух векторов идентифицирующих переменных строгое равенство часто является чрезмерно ограничительным требованием. Алгоритмы увязки записей позволяют

использовать разные метрики сравнения, которые определяют степень, до которой вектора сопоставляемых переменных совпадают. Более того общая парадигма вероятностной увязки записей Феллеги-Сантера включает несколько возможностей, которые могут быть изменены пользователями для удовлетворения их конкретных потребностей. Эти возможности включают метод оценки вероятностей совпадения и правило принятия решения относительно того, какие записи проверять вручную. ESSnet (2008c) и Day (1994) обсуждают методологию для многих аспектов увязки записей.

Выбор сопоставляемых переменных часто включает как ручные, так и автоматические шаги. Аналитики часто имеют предварительные знания о логических совпадающих переменных. Например, номер полиса социального страхования и персональный идентификационный номер являются двумя полезными переменными для сопоставления файлов лиц в США и Нидерландах соответственно. Автоматизированные алгоритмы, часто называемые процедурами профилирования, также были разработаны, чтобы помочь в процессе выбора переменных. Эти автоматизированные процедуры измеряют как корреляцию между переменными в наборах данных, так и качество потенциальных совпадающих переменных в наборе данных. Процедуры профилирования особенно полезны в приложениях с большим количеством потенциальных совпадающих переменных с различной степенью надежности.

Перед выполнением увязки записей форматы сопоставляемых переменных должны быть стандартизированы. В качестве простого примера, два способа записи даты «8 декабря 1952 года» - это «12-8-1952» и «12/8/1952». Алгоритм стандартизации преобразует эти два представления одной и той же даты в формат, который позволит алгоритму сопоставления распознавать два представления как эквивалентные. На практике алгоритмы стандартизации должны работать с более сложными символьными строками, которые представляют такие атрибуты, как номера телефонов, адреса, имена людей, названия предприятий и названия сельскохозяйственных операций. Символьные строки, которые представляют уникальные объекты, могут отличаться, например, из-за тонких различий в использовании заглавных букв, орфографии и пунктуации. Одним из способов стандартизации символьных строк является схема фонетического кодирования. Такие алгоритмы преобразуют строки, которые «звучат одинаково», в уникальный формат символов. Soundex и система идентификации и разведки штата Нью-Йорк (NYSIIS) являются двумя широко используемыми алгоритмами фонетического кодирования.

Сравнение всех пар записей в двух файлах часто неприемлемо с точки зрения объема расчетов. Два метода для уменьшения количества сравнений называются «блокировка» и «сортированная окрестность». При «блокировке» файлы делятся на подгруппы, называемые блоками, и пары записей сравниваются только внутри каждого блока. Для иллюстрации предположим, что два файла, каждый с 5 000 записей, разделены на 10 блоков по 500 записей. Затем количество сравнений уменьшается с $5\,000 \times 5\,000$ до $10(500 \times 500)$. В этом примере количество сравнений, необходимых для незаблокированной структуры, в десять раз превышает количество, требуемое для заблокированной структуры. В процедуре уменьшения размерности методом сортированных окрестностей записи сравниваются, только если они попадают в окно, которое пересекает отсортированные записи.

Поскольку точное совпадение часто слишком ограничительно были, разработаны различные метрики для измерения степени различия двух векторов сопоставляемых переменных. В литературе по увязке записей различные метрики часто называют «функциями сравнения». Альтернативы строгому равенству включают метрику Левенштейна для сравнения двух строк и метрику Джаро-Винклера, которая является специфической для сравнения имен.

При вероятностной увязке записей пользователь может выбирать между различными методами оценки параметров и правилами принятия решений. Например, вероятности совпадения могут быть оценены путем максимизации правдоподобия, которое часто основано на предположении

об условной независимости и реализуется с помощью алгоритма максимизации ожидания (EM алгоритм). Альтернативы методу максимального правдоподобия включают сопоставление на основе частоты или алгоритмические процедуры, которые оценивают модели согласия и несогласия. Алгоритмы увязки записей часто приводят к принятию решений о том, какие записи просматривать вручную. Феллеги и Сантер (Fellegi и Sunter, 1969) доказывают, что конкретное правило принятия решений минимизирует количество пар, которые нужно просмотреть для заданного процента ошибок. Альтернативой процедуре принятия решения Феллеги-Сантера является пороговое правило, согласно которому проверяется любая пара с вероятностью в указанном диапазоне.

ii. Программное обеспечение для микро-интеграции и увязки записей

Проблемы, связанные с объединением разнородных источников данных, носят не только концептуальный характер; они также являются вычислительными. Увязка записей часто включает управление большими объемами данных, поэтому необходимы алгоритмы очистки и стандартизации данных. Для выполнения операций, связанных с очисткой данных и объединением нескольких источников данных, было разработано большое количество программных пакетов. В этом разделе сначала рассматриваются технические возможности существующих программных средств для выполнения шагов, связанных с увязкой записей, которые обсуждались выше. Затем программные инструменты сравниваются по параметрам, которые не связаны строго с техническими возможностями, такими как стоимость, расширяемость и прозрачность.

Обсуждение ниже в основном основано на ESSnet (2008с); тем не менее, оно также содержит идеи из работ Day (1995), Sariyar and Borg (2010) и da Silva *et al.* (2011). ESSnet (2008с) рассматривает несколько пакетов программного обеспечения для увязки записей с точки зрения получения официальной статистики предприятий. Day (1995) рассматривает программное обеспечение для увязки записей с конкретной целью определения наиболее подходящего инструмента для USDA / NASS. В работе Day (1995) содержится обширный и полезный список вопросов и критериев, которые аналитик должен учитывать при выборе подходящего инструмента увязки записей для своих нужд. Конкретные вычислительные инструменты, обсуждаемые в работе Day (1995), могут быть несколько устаревшими. Тем не менее, предлагаемые критерии и вопросы для рассмотрения остаются весьма актуальными. Da Silva *et al.* (2011) рассматривают программное обеспечение для вероятностной увязки записей с целью объединения данных бразильской переписи с данными, полученными из обследования, проведенного в пост-переписной период.

Рассматриваются приведенные ниже пакеты программного обеспечения: первые девять рассмотрены в ESSnet (2008с), а десятый - это пакет R, обсуждаемый в Borg and Sariyar (2010).

1. AutoMatch, разработано в Бюро цензов США, сейчас в компетенции IBM (Herzog *et al.* 2007, гл. 19).
2. Febrl - Freely Extensible Biomedical Record Linkage (Свободно расширяемая увязка биомедицинских записей) разработан в Национальном университете Австралии (FEBRL).
3. Generalized Record Linkage System (GRLS) (Обобщенная система увязки данных), разработан в Статистическом управлении Канады (Herzog *et al.* 2007, гл. 19).
4. LinkageWiz, коммерческое ПО (LINKAGEWIZ).
5. RELAIS, разработан в Национальном институте статистики Италии (ISTAT) (RELAIS).
6. DataFlux, коммерциализирован SAS (DATAFLUX).
7. The Link King, коммерческое ПО (LINKKING).

8. Trillium, коммерческое ПО (TRILLIUM).
9. Link Plus, разработан в Центре контроля и предотвращения заболеваний США (CDC), подразделение онкологии (LINKPLUS).
10. RecordLink, пакет в R разработан Murat Sariyar и Andreas Borg.

Заключения и выводы о возможностях пакетов программного обеспечения следует рассматривать как ориентировочное руководство, а не как стандарт. При принятии решения о том, какие инструменты использовать для конкретного приложения, может быть полезна независимая сравнительная оценка программного обеспечения для определения наиболее подходящего пакета для удовлетворения потребностей конкретного приложения. Приведенный ниже обзор призван обеспечить полезную отправную точку.

Технические возможности

В таблице 4 ниже обобщены технические возможности разных пакетов программного обеспечения. Эти пакеты более подробно рассмотрены в ESSnet (2008с). Столбец «Стандартизация» указывает на возможности обработки и стандартизации, если таковые имеются. В столбце «Профилирование» указано, есть ли в программном обеспечении опции для автоматического профилирования, а «Сокращение пространства» указывает на доступные методы блокировки. Столбец «Правила оценки и принятия решений» содержит информацию о процедуре, используемой для оценки вероятностей совпадения (EM-алгоритм или другое), а также о типе правила, используемого для определения того, считается ли пара записей «совпадением», «несовпадением» или «возможным совпадением». Поскольку все программные пакеты содержат функции сравнения, в столбце «Функции сравнения» указывается масштаб доступных функций сравнения на основе информации, предоставленной в ESSnet (2008с).

ТАБЛИЦА 4. ТЕХНИЧЕСКИЕ ВОЗМОЖНОСТИ ПАКЕТОВ ДЛЯ УВЯЗКИ ЗАПИСЕЙ

Пакет	Стандартизация	Профилирование	Сокращение пространства	Правила оценки и принятия решений	Функции сравнения
AutoMatch	NYSIS, Soundex, другие	Нет	Блокировка	Частотное взвешивание с порогом	Стандартная
Febrl	Скрытые марковские модели и методы на основе базы правил	Нет	Блокировка и сортированная окрестность	Несколько неконтролируемых классификаторов	Большой выбор
GRLS	NYSIS, Soundex, другие	Нет	Блокировка	Модели согласия/несогласия	Стандартная
LinkageWiz	NYSIS, Soundex	Нет	Не определено	Немного подробностей об оценке и методе решения	Стандартная
DataFlux	Инструменты для бизнес-данных	Да	Не определено	Простое, детерминистическое решение	Большой выбор
RELAIS	Нет	Да	Блокировка и сортированная окрестность	Оценка максимального правдоподобия, ручное рассмотрение связей многие-ко-многим	Стандартная
The Link King	Нет	Да	Блокировка	Специальная итеративная процедура оценки и вероятностные и детерминированные правила принятия решений	Большой выбор
Trillium	Широкая	Да	Не определено	Вероятностная не типа Феллеги-Сантер, процедура не определена	Стандартная
Link Plus	Нет	Нет	Блокировка	Максимальное правдоподобие и вероятностное решение	Большой выбор
RecordLink	Нет	Нет	Блокировка	Максимальное правдоподобие с EM алгоритмом и многими классификаторами для правил принятия решений	Стандартная

Удобство использования

В дополнение к сравнению технических возможностей, важное значение имеет понимание удобства использования для выбора соответствующего программного инструмента для увязки записей. Согласно ESSnet (2008с), учитываются следующие показатели удобства использования:

- Стоимость. Программное обеспечение бесплатное или коммерческое? Требуются ли для программного обеспечения лицензии для определенных инструментов управления данными или статистического анализа?
- Специфика домена. Может ли инструмент работать с разными языками или программное обеспечение работает только на английском? Разработан ли инструмент для определенного класса приложений или объектов, таких как бизнес-данные, люди или службы здравоохранения?
- Прозрачность. Хорошо ли документированы процедуры? Может ли аналитик понять, как работают инструменты увязки записей и управления данными?
- Расширяемость. Может ли аналитик изменить и адаптировать процедуры в соответствии со своими конкретными потребностями?
- Выходные отчеты. Удобен ли формат выходных данных? Легко ли использовать связанные файлы и переносить их в другую систему?

Таблица 5 ниже, основанная на ESSnet (2008с), суммирует удобство использования программных инструментов увязки записей.

ТАБЛИЦА 5. ОБЗОР УДОБСТВА ИСПОЛЬЗОВАНИЯ РАЗНЫХ ПАКЕТОВ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Пакет	Стоимость и требования	Специфика домена	Прозрачность	Внедрение
AutoMatch	Коммерческое	Только английский	Богатая документация	Высокое
Febri	Бесплатное	Только английский	Имеется исходный код	Среднее
GRLS	Требует ORACLE	Только английский	Бесплатный, двуязычный учебный курс	Среднее
LinkageWiz	Коммерческое, но дешевое	Английский и французский	Нет точного описания	Среднее
RELAIS	Открытый источник, бесплатное	Нет конкретного домена	Полная доступность исходного кода	Низкое
DataFlux	Требует SASно низкие затраты	Нет конкретного домена	Имеется документация	Высокое
Link King	Бесплатное	Темы здоровья и людей	Хорошо документирован	Среднее
Trillium	Коммерческое	Почти любой язык и страна, но только для маркетинговых приложений	Алгоритмы не точно определены	Среднее
Link Plus	Бесплатное	Онкологические регистры	Отсутствует исходный код, но документация хорошая	Высокое
RecordLink	Бесплатное	Английский и немецкий	Имеются исходный код и документация	Неизвестно, относительно новый пакет

В Таблице 6, основанной на ESSnet (2008с) суммированы сильные и слабые стороны разных программных инструментов.

ТАБЛИЦА 6. ОСНОВНЫЕ СИЛЬНЫЕ И СЛАБЫЕ СТОРОНЫ ПАКЕТОВ ПРОГРАММ ДЛЯ УВЯЗКИ ЗАПИСЕЙ

Пакет	Сильные стороны	Слабые стороны
AutoMatch	Удобная предварительная обработка	Нет оценки степени ошибки. Только английский
Febri	Исходный код, хорошая предварительная обработка и инструменты для сравнения решений из разных алгоритмов увязки записей	Нет профилирования
GRLS	Хорошая предварительная обработка и документация	Только английский, требуется ORACLE
LinkageWiz	Скорость, стандартизация предварительной обработки	Нет профилирования или сокращения пространства, черный ящик
RELAIS	Позволяет заданные пользователем комбинации параметров увязки; адаптируется к широкому спектру ситуаций	Низкое внедрение (новый) относительно мало протестирован
DataFlux	Гибкая предварительная обработка	Детерминистическое решение
Link King	Легко использовать	Не гибкий, нестандартная оценка
Trillium	Удобство и языковая гибкость	Подходит больше для коммерческого применения, чем для официальной статистики, ограниченная документация
Link Plus	Удобно для пользователя бесплатно	Нет предварительной обработки, для онкологических регистров, плохая обработка непоказанных символов во входных данных
RecordLink	Бесплатное, исходный код, много процедур принятия решений, хорошая документация	Нет предварительной обработки, требует стандартизированных входных данных в формате совместимом с пакетом

Для использования в развивающихся странах пакет с характеристиками RELAIS может быть особенно полезен. Программное обеспечение бесплатно, и исходный код полностью доступен. Программное обеспечение позволяет пользователю фиксировать любые комбинации параметров увязки и не является специфическим для конкретной предметной области.

5.1.4. Массовая импутация

Различные регистры и обследования часто содержат разные «переменные ответов». Когда данные увязаны на уровне единиц, наличие нескольких версий взаимосвязанных переменных предоставляет возможность для улучшения качества и расширения. Создание единого полного набора данных, в котором каждая запись появляется один раз, называется «массовой импутацией». Затем данные импутируются для всех записей в результирующую систему регистров. Массовая импутация включает в себя сложные методы моделирования, а вычислительные проблемы возникают в связи с огромным объемом данных (Guigo, 2008).

Например, в обследовании статистики заработной платы и рабочего времени Статистического управления Канады представлены ежемесячные оценки состояния и тенденций для 10 000 заведений. Статистическое управление Канады также имеет доступ к полному файлу форм по денежным переводам для вычетов из заработной платы от таможенной и налоговой служб. Эти административные источники предоставляют данные о численности работающих и валовой месячной заработной плате. Используя эти данные, при помощи регрессионных моделей можно прогнозировать отсутствующие переменные обследования, используя административные переменные в качестве ковариат. Во многих случаях возможна массовая импутация переменных ответа на обследование для всех единиц в административном файле (Grondin and Lavallee, 2001).

5.2. ВИДЫ ИСПОЛЬЗОВАНИЯ ДЛЯ ФОРМИРОВАНИЯ СТАТИСТИЧЕСКОГО ПРОДУКТА (КОСВЕННОЕ ИСПОЛЬЗОВАНИЕ)

В этом разделе показано, как административные данные могут быть использованы для улучшения статистического продукта. Административные данные могут использоваться на всех этапах процесса обследования или переписи, от разработки дизайна выборки до оценки. Административные данные могут использоваться для формирования выборочной совокупности, определения неприемлемых единиц или в качестве вспомогательной информации при построении выборки. Использование административных данных в качестве вспомогательной информации при оценке может повысить эффективность оценок на основе данных обследования или переписи. Хотя эти виды использования более распространены в развитых странах, чем в развивающихся странах, концепции в целом применимы. Чтобы проиллюстрировать, как эти идеи переносятся из развитых в развивающиеся страны, в этом разделе приведены примеры косвенного использования административных данных как в развитых, так и в развивающихся странах.

5.2.1. Построение или улучшение выборочной совокупности

Административные данные часто неразрывно связаны с идентичностью отдельной единицы в целевой совокупности. Многие административные источники создаются в соответствии с процессами, которые определяют конкретные совокупности. Например, данные налогообложения являются результатом процесса сбора налогов и применяются к совокупности налогоплательщиков. Один административный источник данных может использоваться для определения выборочной совокупности.

Лучший подход, чем использование единственного источника данных для непосредственного определения выборочной совокупности, заключается в использовании административных данных из нескольких источников для создания или улучшения основы выборки. Это приводит к улучшению охвата выборочных обследований и переписей (Carfagna и Carfagna, 2010). Примеры из развивающихся и развитых стран приведены ниже:

Примеры из опыта развитых стран

Швеция. Статистическое управление Швеции использует несколько источников налоговой информации для анализа охвата своего регистра предприятий. К этим источникам относятся административные данные, вытекающие из платежей по налогу на добавленную стоимость (НДС), «валовых выплат и предварительного налога на основе отчетов о прибылях и убытках» и «валовой заработной платы, налогов на заработную плату и предварительного налога из ежемесячных налоговых деклараций работодателей» (Berg and Холл, 2007).

Канада. В Канаде существует централизованная статистическая система, в которой Статистическое управление Канады отвечает за сбор и распространение статистической

информации, касающейся демографии, бизнеса, сельского хозяйства и других секторов. Канадский закон о статистике помогает облегчить передачу данных от административных учреждений в Статистическое управление Канады. В Статистическом управлении Канады административные реестры помогают в построении совокупностей для выборки, охватывающих фермы с небольшой площадью земли, такие как фермы по производству кур, яиц, свиней, фруктов и овощей. При отсутствии административных реестров такие фермы трудно отследить (Trant and Whitridge, 2000).

Примеры из опыта развивающихся стран

Индия. Индия и многие другие развивающиеся страны широко используют административные записи и другие формы административных данных для формирования выборочных совокупностей для широкого спектра видов деятельности, таких как: мелкие, средние, крупные или коммерческие и институциональные фермы; данные о сельскохозяйственных животных, такие как отчеты скотобоен и записи о вакцинациях; данные от дилеров или производителей сельскохозяйственного сырья; экспортеров и импортеров. Самая ранняя и, возможно, самая важная форма использования административных записей в индийской статистике касается данных о землепользовании, которые регулярно создаются государственной администрацией по земельному налогу. Эти данные составлены на основании данных о земельных участках в деревнях, которые ведет сельский *патвари* (бухгалтер). Отчеты о землепользовании являются центральными для всего процесса оценки сельскохозяйственного производства в Индии. Они используются в качестве основы выборки, чтобы определить, где следует проводить контрольные уборки урожая. Эти данные также используются в качестве основных исходных статистических данных для оценки производства, которую получают как произведение урожайности, полученной в результате контрольных уборок, и площади под конкретной культурой из записей о землепользовании (Sen, без даты).

5.2.2. Дизайн обследования

Эффективные схемы выборки основаны на информации о структуре рассматриваемой совокупности. Административные источники часто имеют решающее значение в предоставлении внешней информации, необходимой для разработки эффективных выборок. Два примера дизайна выборки, которые используют вспомогательную информацию, - это выборка с вероятностью отбора, пропорциональной размеру (PPS) и стратифицированная выборка. В выборке PPS мера размера определяется для всех единиц в совокупности, а вероятность отбора пропорциональна указанному определению размера. Если показатель размера коррелирует с интересующим ответом, то выборка PPS более эффективна, чем простая случайная выборка. Аналогично, в стратифицированной выборке интересующая совокупность делится на группы, называемые стратами. Если границы страт объясняют изменчивость в совокупности, то стратификация может привести к повышению эффективности. Fuller (2009) и Sarndal, Swensson and Wretman (2005) подробно обсуждают роль вспомогательной информации в дизайне выборки.

Например, в Статистическом управлении Швеции необходимая информация о совокупности для целей дизайна выборки предоставляется из административного источника данных. Использование Статистическим управлением Швеции налоговых данных иллюстрирует роль административных данных в планировании обследований. Статистическое управление Швеции использует налоговые данные для определения страт для обследования акций и активов предприятий. Исследуемая совокупность сильно перекошена, так как на небольшое количество единиц приходится большая доля суммарных значений рассматриваемых переменных. Стратификация совокупности для обследования производится на основе общего объема акций и активов, зарегистрированных в целях налогообложения (Берг и Холл, 2007).

5.2.3. Моделированные калибровочные оценки

Вспомогательные переменные, использующие информацию, «инкапсулированную» в административных данных, часто применяются при оценке так же как и при проектировании выборки. Обоснование, лежащее в основе использования административных данных при оценке, заключается в том, что административные данные могут не соответствовать стандартам, требуемым для статистических данных в некоторых аспектах; однако они имеют нулевую дисперсию выборки и часто коррелируют с величиной, представляющей интерес для обследования. При калибровке весовые коэффициенты для выборочных единиц изменяются таким образом, чтобы соответствующим образом взвешенные суммы вспомогательной переменной были равны контрольному значению из административных данных. Термин «контрольный» используется для обозначения того факта, что оценки подкатегорий должны соответствовать заранее определенной общей величине при объединении, и эта заранее определенная общая величина получена из источников, внешних по отношению к статистическим обследованиям или переписям. Чем сильнее корреляция между переменной, зарегистрированной в административном файле, и переменной обследования, тем больше выигрыш в эффективности от калибровки (Deville, Sarndal and Sautory, 1993). Томсен и Холмой (Thomsen and Holmoy, 1998) приводят примеры и обсуждают использование Статистическим управлением Норвегии административных данных при калибровке.

В некоторых случаях административные данные не предоставляют информацию о точных количествах, а скорее о диапазонах и неравенствах. Например, сумма, полученная из административных записей, которая представляет собой комбинацию более детальных категорий, задает верхнюю границу для итога по любой из участвующих категорий. В таких случаях веса обследования могут быть построены так, чтобы сохранить ограничения неравенства или ограничения диапазона, как они определяются на основании административного источника.

Пример Бюро по управлению земельными ресурсами США иллюстрирует использование таможенных данных для определения ограничения неравенства. Это бюро сотрудничает с USDA/NRCS для получения оценок состояния пастбищных угодий с помощью обследований. Обследование пастбищных угодий 2012 года было направлено на оценку условий обитания шалфейных тетеревов на пастбищных землях бюро в трех областях: приоритетная среда обитания шалфейных тетеревов, экорегионы и зоны подотчетные Западной ассоциации агентств контроля рыбных ресурсов и ресурсов дикой природы. На этапе оценки обследования административные данные о площади пастбищ в 13 западных штатах использовались в качестве контрольных калибровочных элементов при построении весов.

Были использованы следующие административные данные:

- Тематические слои наполнения ГИС, определяющие границы земель, управляемых бюро, в обследуемых штатах от Бюро по управлению земельными ресурсами;
- Тематические слои наполнения ГИС, представляющие совместную работу NRCS и Бюро по управлению земельными ресурсами путем объединения информации о пространственном распределении плотности размножения шалфейных тетеревов с географической базой данных NRCS Common Resource Area;
- Присвоение Агентством по охране окружающей среды США классов экорегионов на основе экорегионов уровня II и уровня III по классификации Омерника (Omernik, 1987); а также
- Тематические слои наполнения ГИС, разграничивающие зоны управления распространением шалфейного тетерева, разработанные Западной ассоциацией агентств контроля рыбных ресурсов и ресурсов дикой природы, отражающие экологические и биологические вопросы и сходства, а не административные границы.

Процедура оценки началась с построения весов для всех точек в выборке, чтобы получить оценки площади пастбищ, управляемых бюро, в каждой комбинации штата, среды обитания и необитания шалфейного тетерева, экорегиона и зоны. Последующее взвешивание включало применение ранжирования и последовательных корректировок отношений для сохранения трех наборов контрольных итогов – страты штатов по типу, экорегионы и зоны Западной ассоциации агентств по рыболовству и дикой природе. В конце калибровки веса, полученные в ходе окончательного анализа, добавляются к административным данным о площадях пастбищных угодий, управляемых бюро, в каждой зоне Западной ассоциации агентств контроля рыбных ресурсов и ресурсов дикой природы.

5.2.4. **Корректировки на неотчеты и импутация**

В обследованиях и переписях обследуемые единицы могут заполнять только часть вопросника или могут отказаться отвечать на вопросы. Если характеристики неответивших систематически отличаются от характеристик ответивших, то оценки, построенные только на полных данных, могут быть смещены из-за того, что изучаемые параметры для совокупности ответивших отличаются. Рассмотрим обследование, предназначенное для получения информации о средних скоростях эрозии пахотных земель. Если фермеры, которые применяют методы охраны земель, имеют более высокие вероятности ответов, то оценки средней эрозии, основанные только на полных ответах на вопросы, могут быть смещены.

Административные данные могут быть доступны как респондентам, так и нереспондентам. Если переменная из административной базы данных наблюдается как для респондентов, так и для нереспондентов, и связана с рассматриваемой переменной из обследования, то вспомогательная информация из административного источника может использоваться для оценки и уменьшения смещения, вызванного отсутствием ответов. Сравнение средних значений вспомогательной переменной для респондентов и нереспондентов может дать представление о природе неотчетов. Если величина, зарегистрированная в административном источнике, коррелирует с интересующими результатами, то административные данные могут использоваться в качестве вспомогательной информации при построении оценок, которые учитывают смещение из-за отсутствия ответа. Два общих метода для корректировки смещения из-за неотчетов – импутация и взвешивание - обсуждаются ниже.

Импутация

Один из механизмов, который может быть использован для корректировки на неотчеты – это импутация пропущенных данных. Импутация особенно полезна в случае неотчетов на отдельные вопросы, то есть когда обследуемая единица заполняет только часть, но не весь вопросник. Если окончательный набор данных сформирован с использованием приемлемого метода импутации, то можно проводить несколько типов статистического анализа о полных данных. Ким и Шао (Kim and Shao, 2013) и Сарндал и Лундстром (Sarndal and Lundstrom, 2005) дают подробное описание теории и методов импутации.

Для описания метода импутации предположим, что данные о векторе изучаемой переменной и вспомогательной переменной (y_i, x_i) собирают в ходе обследования. Пусть A – это множество единиц выборки, обозначаемое $A = \{1, 2, \dots, n\}$ и δ_i показатель ответов, который равен 1, если единица ответила, и 0 - в противном случае. Здесь вспомогательные данные x получают из административного источника, и они имеются как для респондентов, так и для нереспондентов. Для краткости предположим, что механизм ответов характеризуется случайным пропуском данных (*missing-at-random*, MAR), так что:

$$f(y|x, \delta = 1) = f(y|x, \delta = 0) = f(y|x) \quad (6.1)$$

Условие MAR (6.1) означает, что ответы по изучаемой переменной y независимы при вспомогательной переменной x .

При предположении, что принятая процедура выборки является неинформативной (Fuller, 2009), условие MAR выполняется и для совокупности, и для выборок из нее. Поскольку условное распределение y при x и $\delta = 1$ оценивается с использованием данных от респондентов обследования, приемлемые данные для нереспондентов могут быть сформированы на основе оцененной модели импутации, задаваемой $\hat{f}(y|x, \delta = 0)$.

На практике существует два подхода для выполнения импутации. Фракционная импутация (Kim and Shao, 2013) обеспечивает однократно наполненное множество с многократно импутированными значениями по каждой пропущенной единице. Множественная импутация (Rubin, 2004) формирует многократно наполненные множества данных, где каждое наполненное множество данных содержит однократно импутированное значение для каждой пропущенной единицы.

Взвешивание

Альтернативный механизм корректировки на неответы включает изменение весов для учета неответов. Когда окончательная оценка является фиксированной или заранее определенной, взвешивание может быть более эффективным, чем импутация. Калибровочные баллы и мера склонности являются двумя методами определения весов для корректировки на неответы. При калибровке вес определяется таким образом, чтобы среднее значение вспомогательных переменных в единицах выборки было равно среднему значению на основе административных данных. Мера склонности - это оценка вероятности того, что единица i отвечает. Оба метода требуют вспомогательной информации (Lundstrom and Sarndal, 2005), которая может быть получена из административных источников.

Geuzinge, Rooijen и Bakker (2000) описывают использование административных записей при построении весов калибровки для уменьшения смещения, вызванного неответами, в обследованиях домашних хозяйств. В одном приложении административные реестры рабочих мест и пособий по социальному обеспечению использовались для взвешивания респондентов в ходе обследования здоровья, проведенного в Нидерландах в 1995 году. Теория заключалась в том, что люди с более серьезными проблемами со здоровьем и более широким использованием медицинских ресурсов с большей вероятностью будут отвечать на опрос в результате большего интереса к процессам здравоохранения. Опасение статистического агентства заключалось в том, что без корректировки оценка использования медицинских услуг, основанная на полных данных, приведет к завышению реальной стоимости услуг здравоохранения. Оценки дней в больнице, которые включают административные данные, были ниже, чем соответствующие оценки, основанные только на невзвешенных полных данных обследования. Веса также применялись для получения оценок уровней образования. Взвешенные оценки доли лиц, получивших образование выше среднего, были ниже, чем невзвешенные оценки. Считалось, что причиной такого результата было то, что для людей с более высоким уровнем образования была характерна более высокая вероятность ответа, потому что они лучше понимали полезность обследования и больше доверяли правительству.

5.2.5. Моделирование ошибки измерения

Переменные сельскохозяйственного обследования почти всегда подвержены ошибкам измерения. Если ошибка измерения велика, ее игнорирование может привести к искаженным и противоречивым оценкам, которые, в свою очередь, могут привести к ложным выводам. Модели ошибок измерения - это статистические подходы, которые объединяют несколько источников информации в многоуровневой модели для получения единой унифицированной статистики,

связывающей меры неопределенности. В GSARS (2016a) модель ошибки измерения применяется для оценки площади посева кукурузы в Намибии. Используются три источника информации о посевных площадях кукурузы. Одна из оценок получена из ежегодного сельскохозяйственного обследования (AAS), проводимого Намибийским статистическим агентством (NSA). Два получены из министерства сельского хозяйства, водных и лесных ресурсов Намибии. Эти различные источники информации затем объединяются, чтобы получить более точную оценку истинной площади посева. Более подробная информация о модели и процедуре оценки доступна в разделе 5.3.1, GSARS (2016a).

5.2.6. Оценка малых областей

Многие статистические процедуры получения оценок для малых областей или прогнозирования будущих результатов основаны на явных моделях. В случае оценок для малых совокупностей информация о совокупности на уровне малых изучаемых областей очень важна для повышения эффективности оценок. Если цель состоит в прогнозировании или улучшении актуальности оценок, то вспомогательная информация, которая отражает недавний период или изменения во времени, дает возможность сократить средние квадратические ошибки прогнозов. При построении оценок Национальной инвентаризации ресурсов США 1997 года административные данные о транспортировке были использованы для получения оценок малых областей для площади дорог (Nusser, S.M. & Goebel, J.J. 1997b Wang and Fuller, 2003).

В этом разделе кратко рассмотрено, как административные данные можно использовать в качестве вспомогательной информации в классической оценке малых областей. В GSARS (2016b) приведена подробная иллюстрация на основе сельскохозяйственного обследования и административных данных для Намибии.

Рао (2003) подразделяет методы оценки малых совокупностей на модели на уровне единиц и модели на уровне областей. Первые модели были первоначально введены Battese, Harter and Fuller (1988), а последние – Fay and Herriot (1979). Мы рассмотрим модель на уровне области:

$$\widehat{y}_k = y_k + e_k \quad , \quad (6.2)$$

где k обозначает рассматриваемую малую область, \widehat{y}_k – прямая оценка на базе обследования, y_k – неизвестная истинная величина, а $e_k \sim N(0, \sigma_{e,k}^2)$ – ошибка выборки с известной дисперсией $\sigma_{e,k}^2$. Возможно повышение эффективности при построении модели, связывающей истинную величину y_k с вспомогательными контрольными переменными, так что:

$$y_k = x_k' \beta + u_k \quad , \quad (6.3)$$

где x_k – вспомогательная информация, полученная из административных источников, и $u_k \sim N(0, \sigma_u^2)$ – случайный эффект для конкретной области с неизвестной обычной дисперсией σ_u^2 .

При объединении двух моделей (6.2) и (6.3) получим оценку лучшего линейного предиктора для y_k :

$$\widetilde{y}_k = x_k' \hat{\beta} + \widehat{v}_k (\widehat{y}_k - x_k' \hat{\beta}),$$

где $\widehat{v}_k = (\sigma_{e,k}^2 + \sigma_u^2)^{-1} \sigma_u^2$, и $\hat{\beta}$ и σ_u^2 – это оценки, полученные, например, методом максимального правдоподобия или ограниченного максимального правдоподобия.

Отношение средней квадратической ошибки прямой оценки \widehat{y}_k к средней квадратической ошибке предиктора \widetilde{y}_k приблизительно равно \widehat{v}_k . Чем меньше \widehat{v}_k , тем больше повышение

эффективности от использования модели прогнозирования. Другими словами, если истинная величина u_k уже адекватно объяснена административными данными, то может быть получено большое повышение эффективности от включения административных данных в оценку для малых совокупностей.

Battese, Harter и Fuller (1988) используют спутниковые данные в качестве вспомогательной информации при оценке малых областей для посевных площадей и урожая. В их оценке посевных площадей под кукурузой и соевыми бобами в 12 районах штата Айова, вспомогательной информацией в модели для малых площадей было количество пикселей в районе, которые были определены как кукуруза и соевые бобы. Спутниковые данные имели высокую корреляцию с данными обследования, поэтому использование моделей для малых площадей привело к сокращению средней квадратической ошибки по предикторам для малых областей.

Использование моделей оценки малых областей для получения субнациональных оценок площади сельскохозяйственных культур в Намибии изучалось в GSARS (2016b). Доменами были те административные районы Намибии, которые в основном занимались общинным сельским хозяйством. Сельскохозяйственная перепись Намибии 2013/2014 (NCA), выборка общинных сельскохозяйственных угодий, предоставляет данные обследования. В качестве вспомогательной информации использовались данные из Списка вопросов для оценки сельскохозяйственных культур Намибии, системы регулярной отчетности, администрируемой MAWF. В приведенной ниже таблице 7 показаны оценки обследования из Национальной сельскохозяйственной переписи (NCA), оценки, основанные на Списке вопросов для оценки сельскохозяйственных культур MAWF, и предикторы с минимальной среднеквадратической ошибкой (MMSE) на основе модели малых областей. Оценки обычной дисперсии (CV) для обследования и моделей малых областей также предоставляются. Такие оценки CV для предикторов MMSE не могут быть больше, чем оцененные среднеквадратические ошибки для оценок обследования из-за дополнительной информации, содержащейся в модели для малых областей и во вспомогательных данных MAWF. Повышение эффективности в лучшем случае скромное, потому что оценки дисперсий для оценок NCA на региональном уровне относительно низкие.

ТАБЛИЦА 7. ЗНАЧЕНИЯ ДИСПЕРСИИ ДЛЯ ОЦЕНОК, ПОЛУЧЕННЫХ ИЗ НАЦИОНАЛЬНОЙ СЕЛЬСКОХОЗЯЙСТВЕННОЙ ПЕРЕПИСИ (NCA)

Регион	NCA (C.V.)		MAWF	MMSE (C.V.)	
Замбези	15 904	(0.128)	19 384	16 823	(0.122)
Каванго	51 302	(0.090)	21 588	49 999	(0.088)
Омусати	109 673	(0.051)	78 030	109 492	(0.051)
Охангвена	81 337	(0.051)	79 828	81 649	(0.051)
Ошана	40 021	(0.198)	35 100	41 600	(0.177)
Ошикото	68 481	(0.043)	58 080	68 568	(0.042)

Эфиопия

В Африке Эфиопия является одной из стран, которая протестировала использование оценки малых областей. Ежегодные сельскохозяйственные обследования, проводимые Центральным статистическим агентством (CSA), дают оценки посевных площадей только на региональном и зональном уровнях. Из-за небольшого размера выборки, оценки на уровне районов (*werada*) не были доступны. С другой стороны, Министерство сельского хозяйства и развития сельских районов (MoARD) получило оценки площадей на основе данных ARSA. Подход оценки малых областей использовался для разработки оценок районного уровня посевных площадей на основе ежегодных обследований с использованием данных MoARD в качестве вспомогательной переменной (GSARS, 2015c). Для получения дополнительной информации о методологии оценки малых областей см. GSARS (2015b).

5.2.7. Усеченные обследования

В усеченном обследовании весь вопросник или его часть не применяются для части совокупности. Вместо этого необходимая информация получается из внешнего источника. Данные переписи или административные данные часто предоставляют вспомогательную информацию, необходимую для успешного проведения усеченного обследования. Примеры типов внешних источников данных, которые предоставили необходимую информацию для усеченных обследований, включают налоговые данные и информацию от частных корпораций. Во многих странах статистические агентства применяют стратегии проведения усеченных обследований, например, связанных с предприятиями или энергетикой.

Дизайн усеченных обследований опирается на вспомогательную переменную, которая известна для всей совокупности. Общий подход к усеченным обследованиям начинается с упорядочения интересующей совокупности по размеру. Размер единицы часто указывает на важность единицы в получении общей оценки. Например, в обследованиях бизнес-заведений или сельскохозяйственных операций показатель размера может касаться общего количества занятых или площади фермы, соответственно. В типичных случаях применения усеченного отбора единицы с размером, меньше определенной величины, не включаются в выборку. Усеченная выборка может рассматриваться как вид выборки с вероятностью отбора, пропорциональной размеру (PPS), в которой показатель размера для некоторых единиц совокупности равен нулю.

В усеченных обследованиях надежная вспомогательная информация также имеет решающее значение на этапе оценки. Необходимо получить суррогаты для ответов на целевые вопросы, представляющие интерес, для единиц совокупности, которые не были включены в сбор данных. Эти целевые переменные часто получают из административных источников, которые собирают аналогичную или связанную информацию. Понятия, измеренные во внешнем источнике данных, могут отличаться от целевой переменной обследования из-за различий в учетных периодах, охвате или определениях. В таких случаях могут потребоваться модели для калибровки переменных, имеющихся в административном файле, для приведения в соответствие с целевыми концепциями, представляющими интерес для обследования. Ниже приведены некоторые практические примеры:

- В конце 1990-х годов Статистическое управление Канады использовало усеченный дизайн, чтобы снизить нагрузку на респондентов малого бизнеса. Предприятия, которые были слишком малы, чтобы вносить существенный вклад в общую оценку, были помещены в страту «не выбирать», и налоговые данные использовались для получения оценок для этих единиц (Yung, Rancourt and Hidiroglou, 2007).
- Статистическое управление Словении использует усеченное обследование для улучшения своевременности оценок месячных индексов оборота. В совокупности предприятий на три процента самых крупных единиц приходится более 50 процентов общего оборота. Классический вопросник используется для оценки оборота для трех процентов, а налоговые данные используются для оценки ежемесячного оборота в оставшихся единицах совокупности (Seljak, 2007).

5.2.8. Использование административных данных для помощи в сборе данных для обследований и переписей

Административные данные также могут помочь облегчить процессы сбора данных в ходе обследований или переписей. Это особенно верно, когда конкретные характеристики или идентичности единиц выборки неизвестны до тех пор, пока с ними не будет установлен контакт. Один важный случай, особенно для сельскохозяйственных обследований, имеет место в обследованиях, которые частично или полностью основаны на территориальных выборках. В приведенном ниже примере административные источники предоставляют списки с именами и адресами, которые полезны для связи с единицами, первоначально отобранными из территориальной совокупности, а не из списочной совокупности. Карты местности также полезны для помощи сборщикам данных, которые проводят обследования, связанные с сельским хозяйством и природными ресурсами.

В некоторых обследованиях площадей территорий адрес единицы выборки недоступен в совокупности для отбора, и определение местоположения единиц выборки для сбора данных может представлять проблему. В качестве примера можно привести Проект оценки природоохранных мероприятий (SEAP) в США - серию исследований, предназначенных для измерения различных видов потери почвы и питательных веществ на возделываемых полях. Данные FSA использовались для идентификации потенциальных операторов фермы, что сокращало время и усилия, необходимые для поиска оператора, связанного с данной точкой отбора проб. Надежность этих данных варьируется географически. В тех частях страны, где эти данные считаются менее надежными, информация из дополнительных источников была включена в процесс для облегчения контакта с отобранными единицами.

5.2.9. Использование структурных моделей ошибки измерения для объединения многих измерений связанных величин

Распространенным способом формирования единой улучшенной оценки путем объединения многих источников данных является использование структурной модели ошибки измерений.

Этот подход состоит в построении моделей измерения и структурных моделей для описания взаимосвязи между несколькими источниками данных. При заданной формулировке модели параметры могут быть оценены совместно и могут быть построены предикторы рассматриваемой величины. В GARS (2015e) подробно рассмотрена методология, принятая для некоторых схем структур данных.

В данном разделе кратко рассматривается факторный анализ (Fuller, 1987) – подход, который можно использовать при наличии трех источников данных. Более подробно факторный анализ обсуждается в контексте данных о посевных площадях в Намибии в GSARS (2016 и 2016b). Пусть три наблюдаемых источника данных обозначены Y_{i1} , Y_{i2} и X_i , соответственно. Предположим, что три источника данных удовлетворяют условиям

$$\begin{aligned} Y_{i1} &= \beta_{01} + \beta_{11}x_i + e_{i1} \\ Y_{i2} &= \beta_{02} + \beta_{12}x_i + e_{i2} \\ X_i &= x_i + u_i \end{aligned} \tag{6.4}$$

Где x_i – истинное значение изучаемой величины, а (e_{i1}, e_{i2}, u_i) – случайные ошибки. Предполагаемое математическое ожидание и дисперсия распределения истинного значения x_i являются функциями вектора параметров более низкой размерности. Заметим, что предполагается, что X_i – несмещенное измерение x_i , тогда как Y_{i1} и Y_{i2} могут быть смещены. Смещения Y_{i1} и Y_{i2} представлены в параметрах регрессии. Эти искажения могут возникать из-за таких обстоятельств как изменение учетного периода или небольшие различия в формулировке вопроса. Обычно X_i получают с использованием вероятностной выборки обследования, проводимого в контролируемых условиях. Другие два наблюдения могут быть получены из внешних источников данных, таких как административные данные. При предположении, что матрица ковариации вектора $(x_i, e_{i1}, e_{i2}, u_i)$ является диагональной, параметры модели в модели измерения (6.4) можно определить и оценить.

В рамках проведенного пилотного проекта (GSARS, 2016a) была исследована возможность использования структурной модели ошибки измерения для оценки площади посевов с использованием нескольких связанных источников данных в Намибии. Намибия располагает двумя первичными источниками данных о посевных площадях основных сельскохозяйственных культур (кукуруза, сорго и просо). Одним из источников является набор основанных на обследовании оценок, полученных из AAS. Второй источник – административные данные, полученные от MAWF. Оценки обследования AAS рассматриваются как несмещенное измерение (X_i) истинной величины посевной площади. Две оценки MAWF, одна для коммерческого сельского хозяйства и другая для общинного сельского хозяйства, служат смещенными измерениями, представленными Y_{i1} и Y_{i2} в модели ошибки измерения выше. Мы применяем модель ошибки измерения для получения единой оценки посевной площади, которая использует информацию из трех наборов наблюдений. Пример показывает, что дополнительным преимуществом этого подхода является то, что использование модели позволяет оценить меру неопределенности, связанную с предикторами. Подробности доступны в GSARS (2016a), а соответствующий пример приведен в GSARS (2016b).

5.3. ИСПОЛЬЗОВАНИЕ В КАЧЕСТВЕ КОНЕЧНОГО ПРОДУКТА

Административные данные могут использоваться непосредственно в качестве конечного статистического продукта. В этом случае информация, содержащаяся в источнике административных данных, используется непосредственно для основных целей, таких как политика, управление, принятие решений предприятиями и фермерами. В таблице 8 приведены сводные данные об использовании административных данных для статистических целей в 13

африканских странах, которые участвовали в невероятном обследовании, связанном с использованием ADSAS (GSARS, 2015c). Большинство стран используют административные данные напрямую для составления таблиц, формирования и улучшения совокупности для обследования, планирования обследования и прогнозирования урожая. Только две страны используют административные данные в формальных процедурах статистической оценки, таких как калибровка и импутация. Это обследование показывает, что прямое использование, такое как прямое табулирование и прогнозирование урожая, более распространено во многих африканских странах. В этом разделе обсуждаются различные способы использования административных данных непосредственно в качестве конечного статистического продукта.

**ТАБЛИЦА 8. АДМИНИСТРАТИВНОЕ ИСПОЛЬЗОВАНИЕ ADSAS:
ИСПОЛЬЗОВАНИЕ ДЛЯ ФОРМИРОВАНИЯ СТАТИСТИКИ**

Использование в статистике	БУРУНДИ	ЕГИПЕТ	ГАНА	ЛЕСОГО	ЛИБЕРИЯ	ЛИВИЯ	МАВРИТАНИЯ	МАВРИКИЙ	ЮЖНЫЙ СУДАН	ЮЖНАЯ АФРИКА	СУДАН	УГАНДА	ЗАМБИЯ	ВСЕГО/13
Прямое табулирование	0	1	1	0	1	1	1	1	0	0	1	0	1	8
Формирование/улучшение совокупности обследования	0	1	1	1	0	1	0	1	0	1	1	1	1	9
Дизайн обследования	0	1	1	1	0	0	0	0	1	0	1	1	1	7
Калибровка оценок при помощи моделей	0	1	0	0	0	0	0	0	0	0	1	0	0	2
Корректировка на неотчеты (взвешивание)	0	1	0	0	0	0	0	0	0	0	1	1	0	3
Импутация отсутствующих данных обследования	0	1	0	0	0	0	0	0	0	0	1	0	0	2
Оценка малых областей	0	1	0	0	1	0	0	0	0	0	1	1	1	5
Прогнозирование	0	1	1	1	0	0	0	1	0	1	1	1	1	8
Интеграция данных обследования	0	1	0	0	1	0	0	0	0	0	1	0	1	4
Дальнейшая отчетность	1	1	0	1	0	0	0	0	0	0	1	1	1	6

5.3.1. Прямое табулирование

Если административные данные имеют достаточно хорошее качество, они могут использоваться непосредственно для формирования статистического продукта (Brackstone, 1987; Wallgren and Wallgren, 2010). Основываясь на методах, применяемых в Статистическом управлении Канады, Брэкстоун (1987) описывает прямое табулирование как процессы подсчета единиц в файлах, перекрестной классификации по атрибутам и агрегирования количественных переменных, связанных с каждой единицей. Опубликованные оценки событий естественного движения населения, таких как рождение, смерть, иммиграция и эмиграция, часто получают из административных источников (Trant and Whitridge, 2000). Такие события могут касаться людей или предприятий. Примером демографического события в сельском хозяйстве является образование новой сельскохозяйственной операции. Таможенные документы, содержащие информацию об импорте и экспорте, могут служить основой для статистики сельскохозяйственного производства (Trant and Whitridge, 2000). USDA/NASS регулярно публикует информацию об импорте и экспорте сельскохозяйственной продукции (Harris and Clark, 2013).

Прямые публикации административных данных часто основаны на регистре (или системе регистров), определенном Carfagna и Carfagna (2010) и, аналогично, ООН (2011) как систематический набор однозначно идентифицируемых данных на уровне единиц с механизмом обновления. Регистр, который заполняется из нескольких административных источников, может иметь лучший охват и полноту, чем из одного источника. Среди наиболее часто используемых типов регистров - регистры населения и регистры предприятий. Регистры предприятий, как обсуждали Wallgren и Wallgren (2010), могут обеспечить основу для сельскохозяйственной статистики, поскольку работа фермы аналогична работе коммерческого предприятия. Wallgren и Wallgren (2010) обсуждают использование IACS для прямого табулирования. Для некоторых культур, получающих государственные субсидии, данные IACS считаются высоконадежными и, следовательно, непосредственно заносятся в таблицу для получения агрегированных статистических данных о площади, что дает пример прямого использования административных данных для статистического продукта. Объединение базы данных IACS с регистром предприятий обеспечивает дальнейшее улучшения (Wallgren and Wallgren, 2010).

Один из подходов к оценке актуальности прямого табулирования административных данных заключается в сравнении этих данных с оценками, полученными в ходе специальных обследований за определенный период. В Великобритании и Соединенном Королевстве такой анализ выявил разницу только в 4% в поголовье крупного рогатого скота, оцененном по административным данным и по данным обследований с 2003 по 2006 год (см. Таблицу 9 ниже).

ТАБЛИЦА 9. ПОГОЛОВЬЕ СКОТА В ВЕЛИКОБРИТАНИИ И СОЕДИНЕННОМ КОРОЛЕВСТВЕ НА 1 ИЮНЯ ПО ДАННЫМ ОБСЛЕДОВАНИЯ И ИЗ АДМИНИСТРАТИВНЫХ ИСТОЧНИКОВ: 2003 и 2006

Крупный рогатый скот	‘000							
	Великобритания				Соединенное Королевство			
	2003	2004	2005	2006	2003	2004	2005	2006
Июньское обследование	8,823	8,911	8,727	8,635	10,508	10,588	10,392	10,270
Административные данные	9,202	9,300	9,154	8,970	10,946	11,070	10,867	10,657
Разница: обследование – админ.	-380	-390	-428	-335	-438	-482	-475	-386
Процентное различие	-4	-4	-5	-4	-4	-4	-4	-4

Источник: Elliott and McDonnell, 2007

Наличие административных баз данных для целей прямого составления таблиц сельскохозяйственной статистики зависит от административных процессов отдельных стран. За последние 12 лет многие развивающиеся страны создали свои собственные программы сельскохозяйственных субсидий, чтобы конкурировать с ценами в США и Европейском союзе (Clay, 2013). Программы субсидирования в Бразилии, Российской Федерации, Индии, Индонезии и Китае росли быстрее всего (Clay, 2013). Программы страхования фермерских хозяйств представляют собой другой потенциальный источник административных данных в развивающихся странах (Roberts, 2005). В настоящее время страховые программы сосредоточены преимущественно в развитых странах. Однако страхование урожая в развивающихся странах расширяется из-за растущей коммерциализации сельского хозяйства, появления новых страховых продуктов, основанных на погодных индексах, и развития международной торговой политики (Roberts, 2005). Хотя охват таких программ в развивающихся странах в настоящее время может быть недостаточным для прямого составления таблиц, но если они будут продолжать расширяться, то в будущем информацию из этих программ можно усилить использованием в сельскохозяйственной статистики на основе регистров. Ниже приведены некоторые примеры прямого табулирования административных данных:

- NASS публикует административную информацию о забое свиней (Harris and Clark, 2013), полученную в результате проверок, проведенных должностными лицами федерального и регионального уровней. Данные инвентаризации свиней в NASS должны соответствовать опубликованным данным забоя;
- Статистическое управление Канады использует налоговые данные для оценки расходов фермерских хозяйств с целью снижения нагрузки на респондентов. Использование административных данных вместо данных обследований может улучшить качество данных, поскольку считается, что фермеры завышают расходы и занижают объемы продаж в обследованиях (Trant and Whitridge, 2000).

УСЛОВИЯ, ПРИ КОТОРЫХ ПРЯМОЕ ТАБУЛИРОВАНИЕ АДМИНИСТРАТИВНЫХ ДАННЫХ МОЖЕТ БЫТЬ ЦЕЛЕСООБРАЗНЫ ИЛИ НЕОБХОДИМЫ

Если приведенные ниже условия не выполняются, статистическое управление должно рассмотреть вопрос о проведении обследования или переписи.

- Административный источник данных измеряет величину, представляющую интерес для статистического агентства.
- Административный источник данных полностью охватывает целевую совокупность.
- Ошибки отчетности в административном источнике данных незначительны.

5.3.2. Прогнозирование урожая

Административные данные, собранные по различным аспектам погоды (метеорологические данные, информация дистанционного зондирования и т. д.), используются в ряде стран для прогнозирования урожайности или производства в целях обеспечения продовольственной безопасности. Примеры для развивающихся стран приведены ниже:

Замбия: обычные данные о местных тенденциях развития животноводства и растениеводства, собранные специалистами по распространению сельскохозяйственных знаний Министерства сельского хозяйства и кооперативов, используются для составления предварительного прогноза.

Мали: Подразделение раннего предупреждения о голоде (Systeme d'Alerte Precocе, или SAP) собирает и анализирует информацию о прогнозах урожая, спутниковых снимках, ценовых тенденциях и потенциальных угрозах, связанных с климатом или вредителями, для того чтобы заранее предупреждать о надвигающихся продовольственных кризисах и давать рекомендации по действиям для улучшения ситуации (Kelly and Donovan, 2008).

Индия: Схема своевременной отчетности (TRS) преследует основную цель сокращения временного отставания в предоставлении статистических данных о площади основных сельскохозяйственных культур в дополнение к предоставлению совокупности единиц для выборки для выбора участков для выращивания сельскохозяйственных культур, на которых должны проводиться эксперименты по контрольной уборке урожая. В соответствии с TRS, для подготовки предварительных оценок площади под основными культурами, патвари (*patwari*) должны заполнять гирдавари (*girdawari*) в приоритетном порядке для 20-процентной случайной выборки деревень и представлять отчеты о сельскохозяйственных культурах вышестоящим органам власти до установленной даты. Предварительные оценки используются при составлении прогнозов урожая. Это дает правительству Индии предварительные оценки производства, которые имеют решающее значение для принятия различных решений, касающихся ценообразования, распределения, экспорта и импорта.

5.3.3. Обзор прямого использования административных данных в странах

Выборочные обследования и переписи, в которых используются строгие статистические методы, остаются наиболее надежным источником сельскохозяйственных данных. Однако сельскохозяйственные обследования в развивающихся странах проводятся нерегулярно из-за бюджетных ограничений (Pangapanga *et al.*, 2013). Административные записи могут использоваться для заполнения пробелов в сельскохозяйственных данных, если обследования полностью отсутствуют или обеспечивают недостаточный охват ключевых сельскохозяйственных переменных. Например, в Дании, если информация по определенным аспектам имеется в административном регистре, Статистическое управление Дании не включает их в переписной лист или опросный лист обследования (Jensen and Larsen, 2016). Административные данные помогают решать вопросы планирования в существующих государственных системах, и доступны на ежегодной основе (Pangapanga *et al.*, 2013).

Во многих развивающихся странах большая часть данных, производимых и распространяемых через национальные, региональные или глобальные базы данных или публикации, поступают из различных источников из-за отсутствия регулярных статистических обследований или переписей, проводимых странами (Keita and Chin, 2013).

В Таблице 10 ниже показано, что информация из ADSAS важна для разработки, реализации и мониторинга политики в большинстве стран, где были получены ответы на вопросы обследования. Эта информация также используется для поддержки инвестиционных решений, планирования и мониторинга продовольственной безопасности, предоставления информации пользователям для различного использования и для измерения прогресса в реализации международных соглашений и целей.

ТАБЛИЦА 10. ИСПОЛЬЗОВАНИЕ ADSAS В КАЧЕСТВЕ ОКОНЧАТЕЛЬНОЙ СТАТИСТИКИ

Нестатистическое использование	БУРУНДИ	ЕГИПЕТ	ГАНА	ЛЕСОГО	ЛИБЕРИЯ	ЛИВИЯ	МАВРИТАНИЯ	МАВРИКИЙ	ЮЖНЫЙ СУДАН	ЮЖНАЯ АФРИКА	СУДАН	УНАНДА	ЗАМБИЯ	ВСЕГО/13
Формулирование, реализация и мониторинг политики	1	1	1	1	1	0	1	1	1	1	1	1	1	12
Поддержка инвестиционных решений	1	1	1	1	0	1	1	1	0	1	1	1	1	11
Планирование и мониторинг продовольственной безопасности	1	1	1	1	1	0	1	1	0	1	1	0	1	10
Предоставление информации пользователям	1	1	1	1	1	1	1	1	0	0	1	0	1	10
Измерения прогресса в реализации международных соглашений и целей	1	1	1	1	1	0	1	0	0	0	1	1	1	9
Достижение эффективных рынков	0	1	1	0	0	0	0	1	0	0	1	0	0	4

Источник: GSARS (2015с). ПРИМЕЧАНИЕ: 1 = да, 0 = нет

РЕЗЮМЕ

В статистике административные данные могут использоваться различными способами. Первый подход - это прямое использование в качестве сельскохозяйственной статистики при определенных условиях. Кроме того, административные данные могут использоваться для формирования статистического продукта, либо во время разработки дизайна выборки, либо во время обработки и оценки данных. Следовательно, административные данные должны быть интегрированы в систему сельскохозяйственной статистики путем разработки и совершенствования ADSAS.

ГЛАВА 6

ИНТЕГРИРОВАНИЕ

АДМИНИСТРАТИВНЫХ ДАННЫХ В

НАЦИОНАЛЬНЫЕ СИСТЕМЫ

СЕЛЬСКОХОЗЯЙСТВЕННОЙ

СТАТИСТИКИ

Учитывая потенциальное использование административных данных для повышения доступности и качества сельскохозяйственной статистики, эти данные следует принимать во внимание в национальной системе или стратегии, связанной с производством сельскохозяйственной статистики. Краткосрочная цель состоит в том, чтобы разработать и улучшить ADSAS (как обсуждалось в предыдущих главах), чтобы воспользоваться преимуществами административных данных, относящихся к сельскохозяйственному сектору. Затем необходимо будет полностью интегрировать использование данных ADSAS в национальный план сельскохозяйственной статистики, которым могут быть, например, Стратегические планы развития сельскохозяйственной и сельской статистики (SPARS), разработанные и продвигаемые Глобальной стратегией. Долгосрочной целью является развитие системы сельскохозяйственной статистики на основе регистров посредством разработки или совершенствования ключевых регистров в стране.

6.1. ИНТЕГРАЦИЯ ADSAS В SPARS

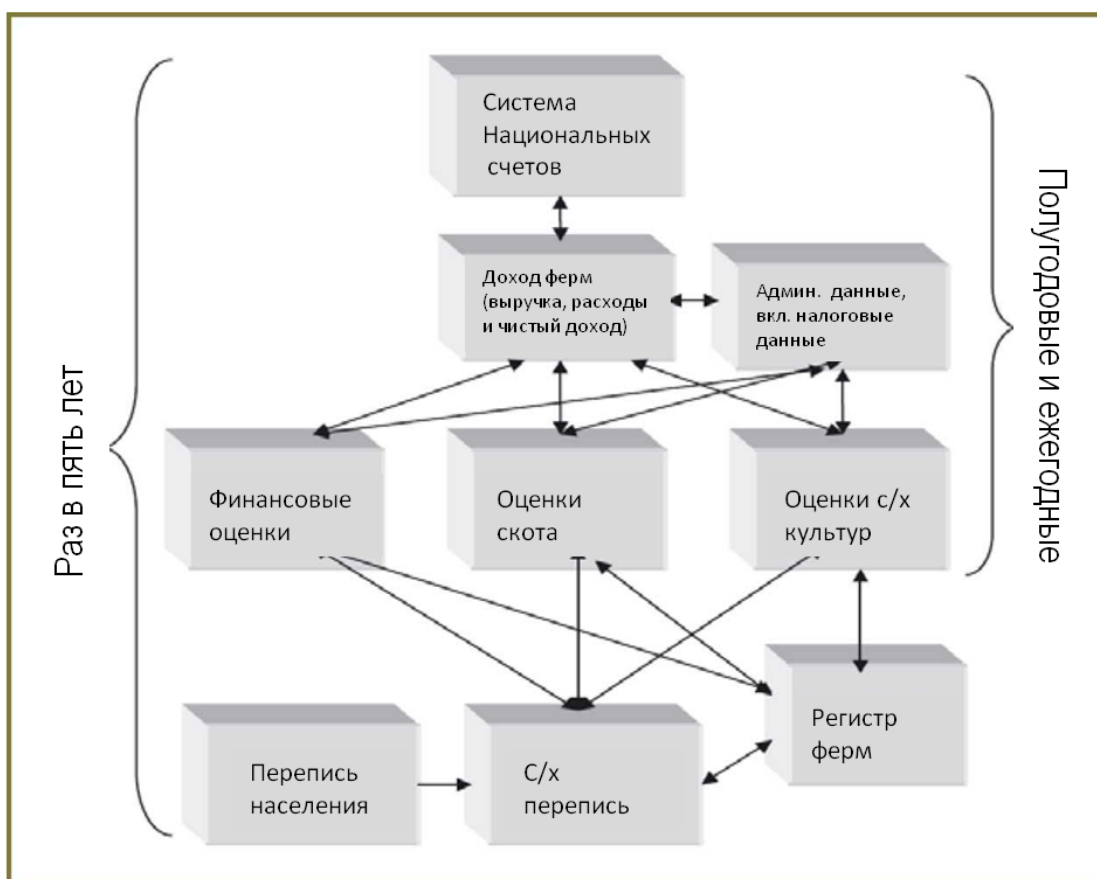
Для каждой страны крайне важно разработать четкую стратегию в отношении производства сельскохозяйственной статистики, которая должна учитывать совершенствование и использование административных данных. Глобальная стратегия рекомендует странам разработать SPARS (Стратегические планы развития сельскохозяйственной и сельской статистики) с учетом административных данных и разработать стандартную методологию для подготовки этих планов (GSARS, 2014). SPARS обеспечивают основу для определения сильных сторон и приоритетов политики, а также соответствующих потребностей в данных, критических пробелов, недостатков, дублирования и несоответствий. Они должны охватывать весь сельскохозяйственный и сельский сектор, включая весь сбор, анализ, распространение и использование данных переписей, обследований и административных систем (GSARS, 2014). Методология разработки SPARS включает оценку статистических результатов из существующих источников сельскохозяйственных данных, включая системы административной отчетности.

Методология рекомендует, чтобы SPARS были реалистичными и прагматичными в отношении ресурсов. Это подразумевает, например, расстановку приоритетов, последовательность и экономическую эффективность, учет альтернативных способов сбора данных, таких как сбор из административных источников и выборочных обследований. Надлежащая интеграция

административных данных в сельскохозяйственную статистическую систему обеспечит значительную экономию средств, как описано в GSARS (2015b).

Пример интеграции ADSAS в национальную систему сельскохозяйственной статистики приведена на рис. 4 ниже.

РИСУНОК 4. СХЕМА ОРГАНИЗАЦИИ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИКИ В СТАТИСТИЧЕСКОМ УПРАВЛЕНИИ КАНАДЫ



Source: Dion, Chartrand, and Murray (2010).

6.2. ДОЛГОСРОЧНАЯ ПЕРСПЕКТИВА ИНТЕГРАЦИИ АДМИНИСТРАТИВНЫХ ДАННЫХ В СТАТИСТИЧЕСКУЮ СИСТЕМУ

В Глобальной стратегии совершенствования сельскохозяйственной и сельской статистики (Всемирный банк, ФАО и ООН, 2011 г.) обсуждается интеграция сельского хозяйства в национальную статистическую систему для улучшения сельскохозяйственной статистики. Долгосрочной перспективой этой интеграции может быть система сельскохозяйственной статистики на основе регистров. Wallgren и Wallgren (2017) описывают, как производятся статистические данные о сельском хозяйстве и сельском населении в стране, где производство статистики основано на регистрах, и обсуждают модернизацию национальной статистической системы из традиционной системы переписей в систему на основе регистров. Здесь предлагается долгосрочная стратегия, которая тесно следует руководящим принципам, предложенным Wallgren и Wallgren (2017), для создания интегрированной системы статистических регистров с

использованием административных источников данных. Предполагается составить план создания интегрированного статистического регистра ферм, который можно регулярно обновлять на основе нескольких источников административных данных.

6.2.1. Модернизация интегрированной системе на основе регистров

Микроданные с идентификаторами

Административные регистры содержат *идентификаторы*. Идентификационные номера играют важную роль в построении интегрированной системы на основе регистров: они способны оценивать качество данных и соединять несколько источников данных, используя такие методы, как детерминированная увязка записей. Поэтому отправной точкой для интегрированной системы является обследование или перепись на основе регистров в сочетании с административными данными, увязанными с использованием идентификаторов.

Совершенствование административной системы

Административные данные должны быть такого же качества как данные обследования. Ключевые переменные, такие как идентификационные номера, регистрация рождений и смертей и миграции, так же важны для увязки данных из разных источников.

Защита конфиденциальности

Идентификационная информация должна быть анонимной для физических и юридических лиц. Обеспечение конфиденциальности имеет решающее значение, если данные должны свободно увязываться с другими наборами административных данных без риска раскрытия личной информации.

Централизация, сотрудничество и законодательство

Другим ключевым условием является построение централизованной статистической системы. Wallgren и Wallgren (2007) рекомендовали, чтобы *«национальный статистический институт в стране отвечал за все регистры, которые заменяют перепись населения и жилищного фонда, и за все регистры, которые используются для национальных счетов»*. Кроме того, для эффективной работы административных систем необходимо сотрудничество между национальными статистическими институтами и другими организациями.

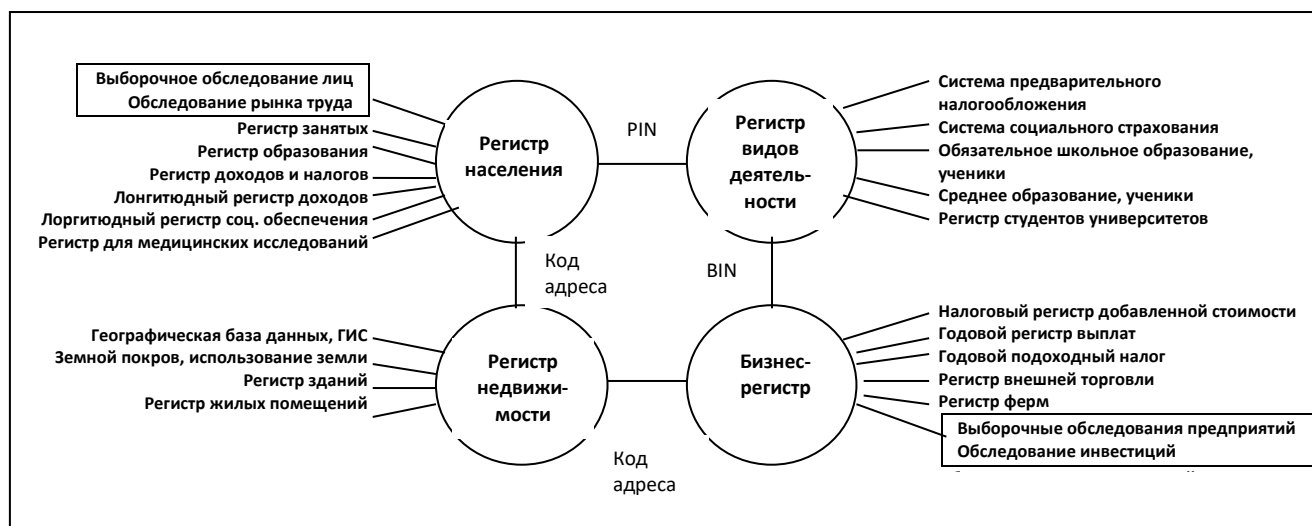
Оценка качества

Как обсуждалось в главе 5, на административные данные часто влияют методология, выборка и другие вопросы качества данных. В этом отношении ситуация с регистрами схожа: на данные часто влияет охват ошибки измерений. Таким образом, административные данные должны быть оценены до их применения с использованием различных показателей качества. Систематическая проверка качества входных данных обсуждается в Wallgren и Wallgren (2014).

6.2.2. Создание интегрированной, основанной на регистре, системы для сельскохозяйственной статистики

Необходимым условием модернизации является наличие интегрированной статистической системы, на основе которой может быть создан согласованный регистр. На рисунке 5 показана общая система статистического производства с основными регистрами, другими статистическими регистрами и выборочными обследованиями. Термин «интегрированная система» используется, чтобы подчеркнуть, что совокупности и переменные непротиворечивы, а оценки согласованы.

РИСУНОК 5. ПРОИЗВОДСТВЕННАЯ СИСТЕМА НА БАЗЕ РЕГИСТРОВ, СИСТЕМА РЕГИСТРОВ И ВЫБОРОЧНЫЕ ОБСЛЕДОВАНИЯ.



PIN – идентификационный номер физических лиц BIN - идентификационный номер предприятий

Создание новых статистических систем зависит от построения системы регистрации. Процедуры по созданию систем, предложенные Wallgren и Wallgren (2007) приведены ниже.

- Шаг 1.** Создание национальной системы регистрации с хорошими идентификационными номерами. Эти персональные идентификационные номера должны использоваться в разных административных системах.
- Шаг 2:** Разработка статистического регистра населения, который может быть основан на административном регистре населения и дополнен другими источниками для улучшения охвата резидентских адресов..
- Шаг 3 :** Разработка регистра объектов недвижимости или кадастра, регистра предприятий и регистра ферм. Регистр занятых и регистр образования так же существенны в системе, основанной на регистрах.

План, предложенный Wallgren и Wallgren's plan (2007), считается долгосрочным, поскольку создание необходимых условий и шаги по построению требуют значительного времени и затрат. Небольшой масштаб сельскохозяйственного производства во многих развивающихся странах делает построение, не говоря о ведении, регистра ферм очень сложным (в лучшем случае). Наиболее практичным вариантом могло бы быть построение регистров для институциональных и крупных ферм. К сожалению, во многих развивающихся странах на их счет приходится только небольшая часть общего объема производства. Другим недостатком этой стратегии является то, что некоторые административные данные, такие как экспертные оценки и оценки на глаз, не обрабатываются в интегрированной системе на основе регистров. Однако если имеются финансовые ресурсы и институциональная поддержка, система на основе регистров является идеальным методом для сбора сельскохозяйственных данных.

Ряд развивающихся стран, особенно в Азии, разработали и даже оцифровали свои земельные регистры. Это улучшает качество сельскохозяйственных административных данных и является хорошим примером для тех развивающихся стран, которые не имеют земельных регистров.

РЕЗЮМЕ

В этой последней главе обсуждается важность интеграции административных данных в системы сельскохозяйственной статистики. Рекомендуется учитывать улучшение и использование этих данных в стратегических планах по производству сельскохозяйственной статистики. Подчеркивается, что SPARS, разработанные и рекомендованные Глобальной стратегией, должны учитывать улучшение и использование административных данных, как подробно обсуждалось в предыдущих главах этой публикации. Наконец, стратегия и методы для развития системы сельскохозяйственной статистики на основе регистров обсуждаются в качестве долгосрочной перспективы.

Литература

Abaye, A. 2010. *The 2003 Pastoral Areas Livestock Census and the Data Needs in the National Strategy for the Development of Statistics (NSDS) in Ethiopia*. Paper presented at the Fifth International Conference on Agricultural Statistics, Kampala. 12-15 October 2010. Available at: www.fao.org/fileadmin/templates/ess/documents/meetings_and_workshops/ICAS5/PDF/ICASV_1.1_008_Paper_Aberash.pdf. Last accessed 12 April 2017.

Administrative Data Liaison Service (ADLS). 2017. *Administrative data introduction*. Available at: <http://www.adls.ac.uk/adls-resources/guidance/introduction/>. Last accessed 12 April 2017.

African Development Bank (AfDB), Partnership in Statistics for Development in the 21st Century (PARIS21) & Intersect. 2007. *Mainstreaming Sectoral Statistical System: A Guide to Planning a Coordinated National Statistical System*. AfDB Publication: Tunis.

Asian Development Bank (ADB). 2010. *Administrative Data Sources for Compiling the Millennium Development Goals and Related Indicators: A reference handbook on using data from education, health, and vital registrations systems featuring practices and experiences from selected countries*. ADB Publication: Manila.

Bakker, F.M.B. 2012. Estimating the validity of administrative variables. *Statistica Neerlandica*, 66: 8-17.

Barboza, W.J. & Harris, M. 2009. *Utilizing an Alternative Sampling Frame to Produce Agricultural Survey Indications*. Paper prepared for the 2009 Research Conference of the Federal Committee on Statistical Methodology. NASS/USDA Research and Development Division. Available at: http://www.nass.usda.gov/Education_and_Outreach/Reports,_Presentations_and_Conferences/reports/conferences/FCSM/fcsm2009_paper.pdf. Last accessed 12 April 2017.

Battese, G.E., Harter, R.M. & Fuller, W.A. 1988. An error-components model for prediction of county crop areas using survey and satellite data. *Journal of the American Statistical Association*, 83: 28-36.

Beckler, D.G. 2013. *Administrative Data Used by the National Agricultural Statistics Service*. Proceedings of the Sixth International Conference on Agricultural Statistics, 23-25 October 2013. Rio de Janeiro, Brazil.

Benedetti, R., Espa, G. & Lafratta, G. 2005. *Tree-based approach to forming strata in multipurpose business survey*. Discussion paper No. 5. Publication of the Università Degli Studi Di Trento - Dipartimento di Economia: Trento, Italy.

Berg, S., Hall, J., Farnstrand, J. & Norlander, M. 2007. *Administrative data in the Swedish SBS*. Statistics Sweden Publication: Stockholm.

Berka, C., Humer, S., Moser, M., Lenk, M., Rechta, H. & Schwerer, E. 2012. Combination of evidence from multiple administrative data sources: quality assessment of the Austrian register-based Census 2011. *Statistica Neerlandica*, 66: 18-33.

- Biemer, P., Trewin, D., Bergdahl, H. & Japac, L.** 2014. A System for Managing the Quality of Official Statistics. *Journal of Official Statistics*, 30: 381-115.
- Bins, B.O. & Dale, P.F.** 1995. *Cadastral surveys and records of rights in land*. FAO Land Tenure Studies. FAO Publication: Rome.
- Borg, A. & Seriyar, M.** 2010. The Record Linkage Package: Detecting Errors in Data. *The R Journal*, 2: 61-67.
- Brackstone, G.** 1987. Statistical Issues of Administrative Data: Issues and Challenges. *Survey Methodology*, 13(1): 29-13.
- Bycroft, C.** 2010. *A register-based census: What is the potential for New Zealand?* Statistics New Zealand Publication: Wellington.
- Carfagna, E. & Carfagna, A.** 2010. *Alternative sampling frames and administrative data; which is the best data source for agricultural statistics?* In Benedetti, R., Bee, B., Espa, G. & Piersimoni, F. (eds), *Agricultural Survey Methods*. John Wiley & Sons: Chichester, United Kingdom.
- Chambers, R.** 2009. *Regression analysis of probability-linked data*. Official Statistics Research Series, vol. 4, Statistics New Zealand.
- Clay, J.** 2013. "Are agricultural subsidies causing more harm than good?" *The Guardian*, 8 August 2013. Available at: <http://www.theguardian.com/sustainable-business/agricultural-subsidies-reform-government-support>. Last accessed 12 April 2017.
- Daas, P.J.H. & Fonville, T.C.** 2007. *Quality control of Dutch Administrative Registers: An inventory of quality aspects*. Paper prepared for the Seminar on Registers in Statistics - methodology and quality, Helsinki, 21-23, May 2001.
- Daas, P.J.H., Ossen, S.J., Tennekes, M. & Burger, J.** 2012a. *Evaluation and visualisation of the quality of administrative sources used for statistics*. Paper prepared for the European Conference on Quality in Official Statistics, 29 May-1 June 2012, Athens.
- Daas, P.J.H., Puts, M.J., Buelens, B. & Van den Hurk, P.A.M.** 2012b. *Big Data and Official Statistics*. Available at http://www.pietdaas.nl/beta/pubs/pubs/NTTS2013_BigData.pdf. Last accessed 12 April 2017.
- Da Silva, D.A., Romeo, S.M.O., Soares, T.S. & Xavier, V.L.** 2011. Study of Record Linkage Software for the 2010 Brazilian Post Enumeration Survey. Proceedings of the 58th World Statistical Congress, Dublin, pp. 1056-1063.
- Day, C.** 1995. Record Linkage 1: Evaluation of Commercially Available Record Linkage Software for Use in NASS. United States Department of Agriculture, National Agricultural Statistics Service, Research Division. USDA Publication: Washington, D.C.
- Deville, J.C., Sarndal, C.E., & Sautory, O.** 1993. Generalized raking procedures in survey sampling. *Journal of the American Statistical Association*, 88: 1013-1020.
- Dion, M.** 2007. *Metadata: an integral part of Statistics Canada's data quality framework*. Paper prepared for the Fourth International Conference on Agricultural Statistics (ICAS-IV), 22-21 October 2007, Beijing.

Dion, M., Chartrand, D. & Murray, P. 2010. *Statistics Canada's Quality Assurance Framework applied to agricultural Statistics*. In: R. Benedetti, M., Bee, Espa, R. & Piersimoni, F. (eds) *Agricultural Survey Methods*. John Wiley & Sons: Chichester, United Kingdom.

Elliott, M. & McDonnell, P. 2007. *Request from the United Kingdom to Use the Bovine Registers of Great Britain and Northern Ireland in Replacement of Statistical Surveys*. Publication of the Department for Environment Food and Rural Affairs Surveys, Statistics and Food Economics Division: London.

European Communities. 2003. *Methodology of animal statistics*. Publication of the Office for Official Publications of the European Communities: Luxembourg.

Fay, R. & Herriot, R. 1979. Estimation of Income from Small Places: An Application of James-Stein Procedures to Census Data. *Journal of the American Statistical Association*, 74(366): 269-277.

Fellegi, I.P. & Sunter, A.B. 1969. A theory for record linkage. *Journal of the American Statistical Association*, 64(328): 1183-1210.

Fuller, W.A. 2009. *Sampling Statistics*. John Wiley and Sons: New Jersey, USA.

Galmes, M. 2013. *Integrating expert opinion in agricultural statistics*. Paper presented at the *Sixth International Conference on Agricultural Statistics*, Rio de Janeiro, Brazil, 23-25 October 2013.

Geuzinge, L., van Rooijen, J. & Bakker, B.F.M. 2000. *Integrating administrative registers and household surveys*. Statistics Netherlands Publication: Voorburg/Heerlen.

Goebel, J.J. 2009. *Statistical Methodology for the NRI-CEAP Cropland Survey*. USDA/NRCS Report. USDA Publication: Washington, D.C.

Goel, B. 2002. Agricultural Surveys in India - Some Thoughts. *Journal of the Indian Society of Agricultural*

Statistics, 55: 32-46.

Global Strategy to improve Rural and Agricultural Statistics (GSARS). 2014. *SPARS: Strategic Plan for Agriculture and Rural Statistics*. GSARS Technical Report: Rome.

2015a. *Reviewing the Relevant Literature and Studies on the Quality and Use of Administrative Sources for Agricultural Data*. Technical Report 1. GSARS Technical Report: Rome.

2015b. *The Role of Administrative Data in Developed Countries: Experiences and Ongoing Research*. Technical Report 2. GSARS Technical Report: Rome.

2015c. *Critical Analysis of Agricultural Administrative Sources Being Currently Used by Developing Countries*. Technical Report 3. GSARS Technical Report: Rome.

2015d. *Analysis of Agricultural Administrative Data Gaps and Ways of Improving the Quality and Use of Data Sources for Agricultural Statistics*. Technical Report 4. GSARS Technical Report: Rome.

2015e. *Strategy and Methodology for Improving the Use of Administrative Data - A Protocol for in-Country Testing*. Technical Report 5. GSARS Technical Report: Rome.

2015f. *Expert Meeting on Improving Methods for Estimating Crop Area, Yield, and Production under Mixed and Continuous Cropping, 15-16 April 2015*. FAO Headquarters, Rome.

2016a. *Findings from the Field in-Country Testing*. Technical Report 6. GSARS Technical Report: Rome.

2016b. *Strategy and Methodology for the Improvement of the Collection and Management of Administrative Data in an Agricultural Statistics System*. Technical Report 7. GSARS Technical Report: Rome.

2017. *Improving the methodology for using administrative data in an agricultural statistics system. Final Report*. Technical Report n.24. Global Strategy Technical Report: Rome

Grondin, C. & Lavallee, P. 2001. *Survey of Employment, Payroll and Hours: An Update*. Statistics Canada Internal Document: Ottawa. Available at: <http://www.oecd.org/std/30036134.pdf>. Last accessed on 12 April 2017.

Guigo, M. 2008. *Micro-integration of different sources: Introduction and preliminary issues*. In *Proceedings of the ESSnet Statistical Methodological 98 Project on Integration of Survey and Administrative Data: Report of WP2. Recommendations on the use of methodologies on the integration of surveys and administrative data*. Available at: <http://cenex-isad.istat.it/>. Last accessed on 12 April 2017.

Hamer, H. 2013. *Weekly Crop Progress and Condition at USDA-NASS*. Proceedings of the *Sixth International Conference on Agricultural Statistics*, Rio de Janeiro, Brazil, 23-25 October 2013.

Harris, J.M. & Clark, C.Z.F. 2013. *Strengthening Methodological Architecture with Multiple Frames and Data Sources*. Proceedings of 59th ISI World Statistics Congress, 25-30 August 2013, Hong Kong SAR.

Herzog T.N., Scheuren F.J. & Winkler, W.E. 2007. *Data Quality and Record Linkage Techniques*. Springer Science+Business Media: New York, USA.

Hidiroglou, M.A., Rancourt, E. & Yung, W. 2007. *Administrative Data in Statistics Canada's Business Surveys: The Present and the Future*. Statistics Canada Internal Report.

Holt, D. 2007. The Official Statistics Olympic Challenge: Wider, Deeper, Quicker, Better, Cheaper. *The American Statistician*, 61(1): 1-8.

Republic of India. 2013. *Report of the Committee on Statistics of Agriculture and Allied Sectors*. Publication of the National Statistical Commission Ministry of Statistics and Program Implementation: New Delhi.

Government of India Planning Commission. 2001. *Report of the Working Group on Agricultural Statistics*. Government of India Planning Commission Publication: New Delhi.

ISTAT, CBS, GUS, INE, SSB, SFSO & EUROSTAT, Statistical Methodology Project on Integration of Survey and Administrative Data (ESSnet-ISAD) (eds. 2008a. *Report on WPI: State of the art on Statistical Methodologies for Data Integration*. Available at: www.istat.it/it/files/2013/12/FinalReport_WP1.pdf. Last accessed on 12 April 2017.

2008b. *Report of WP2: Recommendations on the Use of Methodologies for the Integration of Surveys and Administrative Data*. Available at: <http://cenex-isad.istat.it/>. Last accessed on 12 April 2017

2008c. *Report of WP3: Software Tools for Integration Methodologies*. Available at: <http://cenex-isad.istat.it/>. Last accessed on 12 April 2017.

Isad, E. 2008. *ESSnet Statistical Methodology Project on Integration of Survey and Administrative Data Report of WP3: Software tools for integration methodologies*. Available at: <http://cenex-isad.istat.it/>. Last accessed on 12 April 2017.

Iwig, W., Berning, M., Marck, P. & Prell, M. 2013. *Data Quality Assessment Tool for Administrative Data*. Available at: <http://www.bls.gov/osmr/datatool.pdf>. Last accessed on 12 April 2017.

Jensen, P.V. & Larsen, K. 2016. *Register based statistics - Agriculture*. Component C Infrastructure for Agricultural Statistics. Activity C.1.a Data sources for agricultural statistics and farms frame. Available at: <http://www.dst.dk/ext/6046544347/0/israel2016/Annex-C1a-10-Register-based-statistics-Agriculture--pdf>. Last accessed on 7 October 2017.

Keita, N. & Chin, N. 2013. *The place of 'assessment' in current agricultural statistics for developing countries: making best use of available information for timely crop production estimates in the absence of a system of agricultural sample surveys*. Paper presented at the *Sixth International Conference on Agricultural Statistics*, Rio de Janeiro, Brazil, 23-25 October 2013.

Kelly, V. & Donovan C. 2008. *Agricultural Statistics in Sub-Saharan Africa: Differences in Institutional Arrangements and their Impact on Agricultural Statistics Systems: A Synthesis of Four Country Studies*. Michigan State University International Development Working Paper No. 95. Michigan State University Publication: East Lansing, MI, USA.

Kim, G. & Chambers, R. 2012. Regression Analysis under Probabilistic Multi-Linkage. *Statistica Neerlandica*, 66(1): 64-70.

Kim, J.K. & Shao, J. 2014. *Statistical Methods for Handling Incomplete Data*. CRC Press: Boca Raton, FL, USA.

Kizito, A. 2011. *The structure, conduct, and performance of agricultural market information systems in Sub-Saharan Africa*. Michigan State University, East Lansing, MI, USA (Ph.D. Dissertation).

Laitila T., Wallgren A. & Wallgren, B. 2011. *Quality Assessment of Administrative Data Research and Development*. Statistics Sweden Publication: Stockholm.

Larsen, M.D. & Rubin, D.B. 2001. Iterative automated record linkage using mixture models. *Journal of the American Statistical Association*, 96: 32-41.

Republic of Malawi. National Statistical Office. 2012b. *National Statistical Office Strategic Plan 2012-2016*. Available at: www.paris21.org/sites/default/files/MALAWI_NSO_SP_2012-2016.pdf. Last accessed on 12 April 2017. National Statistical Office Publication: Lilongwe.

Maligalig, D.S. 2017. *Administrative Reporting Systems for Agriculture in Asia*. Paper presented at the 61st ISI World Statistics Congress, 16-21 July 2017, Marrakech, Morocco.

Namibia Statistics Agency. 2014. *Press Release Update on the Census of Agriculture*. Available at: <http://cms.my.na/assets/documents/p19dmo8lrf15ok8t71u3t1j7k7q1.pdf>.

Nordbotten, S. 2008. *The Use of Administrative data in Official Statistics - Past, Present, and Future - with Special Reference to the Nordic Countries*. In Carlson, M., Nyquist, H. & Villani, M. (eds), *Official*

Statistics in Honour of Daniel Thorburn (pp. 205-223). Available at: www.nordbotten.com. Last accessed on 12 April 2017.

Nusser, S.M. & Goebel, J.J. 1997. The National Resources Inventory: a long-term multi-resource monitoring programme. *Environmental and Ecological Statistics*, 4(3): 181-204.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. Map (scale 1:7,500,000). *Annals of the Association of American Geographers*, 77: 118-125.

Pangapanga, P., Kanyanda, S., Kussein, G. & Tembo, L. 2013. *Economic constraints in agricultural statistics: could administrative data complement agricultural surveys in Malawi National Statistical Systems?* Paper presented at the *Sixth International Conference on Agricultural Statistics*, Rio de Janeiro, Brazil, 23-25 October 2013.

Prell, M., Bradsher-Fredrick, H., Comisarow, C., Cornman, S., Cox, C., Denbaly, M., Wilkie Martinez, R., Sabol, W. & Vile, M. 2009. *Profiles of Success of Statistical Uses of Administrative Data*. Report of a subcommittee of the Federal Committee on Statistical Methodology, U.S. Office of Management and Budget. U.S. Office of Management and Budget Publication: Washington, D.C. Available at: <http://www.bls.gov/osmr/fcsm.pdf>.

Rao, J. 2003. *Small Area Estimation*. John Wiley & Sons: Chichester, UK.

Roberts, R.A.J. 2005. *Insurance crops in developing countries*. FAO Agricultural Services Bulletin. Available at: <ftp://ftp.fao.org/docrep/fao/008/y5996e/y5996e00.pdf>.

Sariyar, M. & Borg, A. (2010). The Record Linkage Package: Detecting Errors in Data. *The R Journal*, 2: 61-67

Sarndal, C.E., Swensson, B. & Wretman, J. 2005. *Model Assisted Survey Sampling*. Springer-Verlag: New York, USA.

Sarndal, C.E. & Lundstrom, S. 2005. *Estimation in Surveys with Nonresponse*. John Wiley & Sons: Chichester, England, UK.

Seljak, R. 2007. Use of the tax data for the purposes of the short-term statistics. Seminar on Registers in Statistics - methodology and quality, Helsinki.

Sen, P. Undated. *Challenges of Using Administrative Data for Statistical Purposes: India Country Paper*. Available at: <https://unstats.un.org/unsd/dnss/docViewer.aspx?docID=2426>.

Smith, J., Beaulieu, M., Kumar, E. & O'Neill, L. 2013. *Recent developments in the production of agriculture statistics*. Paper prepared for the *Sixth International Conference on Agricultural Statistics*, 23-25 October 2013. Rio de Janeiro, Brazil.

Srivastava, A. Undated. *Agricultural Statistics System in India*. Available at: http://www.iasri.res.in/ebook/TEFCPI_samplng/AGRICULTURAL%20STATISTICS%20SYSTEM%20IN%20INDIA.pdf. Last accessed on 12 April 2017.

United Republic of Tanzania. 2011. *Monitoring and Evaluation Framework; Revised Final Draft, March 2011*. Publication of the Agricultural Sector Development Programme (ASDP) M&E Thematic Working Group: Dodoma.

Bureau of Statistics. 2012. *Assessment of the Improved Agricultural Routine Data System (ARDS)*; December 2012. Bureau of Statistics Publication: Dodoma.

Bureau of Statistics. 2012. *Assessment of the Agricultural Routine Data System in Tanzania*. JICA, December, 2012. Available at: www.resakss.org/2014conference/docs/Tanzania_JSR_Assessment.pdf. Last accessed on 12 April 2017.

Thomsen, I. & Holmøy, A.M.K. 1998. Combining data from surveys and administrative record systems. The Norwegian experience. *International Statistical Review*, 66, 201-221.

Trant, M. 2011. *Mission Report, Master Plan Project for Agricultural Statistics*, 17th-29th April 2011. Ministry of Agriculture/INE Publication: Maputo.

Trant, M. & Whitbridge, P. 2000. *Integration of Administrative Data with Survey and Census Data*. Agriculture and Rural Working Paper Series, 42. Statistics Canada Publication: Ottawa. Available at: <http://www5.statcan.gc.ca/olc-cel/olc.action?ObjId=21-601-M1999042&ObjType=46&lang=en>. 12 April 2017.

Turtoi, C. Akyildirim, O. & Petkov, P. 2012. Statistical Farm Register in the EU Acceding Countries - A Conceptual Approach. *Economics of Agriculture* 1/2012.

Uganda Bureau of Statistics (UBOS). 2007. *The Development of the Agricultural Sector Strategic Plan for Statistics: A Data Collection Plan for Agricultural Statistics in Uganda*. Final Report to the Uganda Bureau of Statistics by the National Consultant: February 2007. UBOS Publication: Kampala.

United Nations (UN). 2011. *Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices*. Available at: <http://unstats.un.org/unsd/EconStatKB/KnowledgebaseArticle10349.aspx>. Last accessed on 12 April 2017.

United Nations Economic and Social Council. 2007. Seminar on Increasing the Efficiency and Productivity of Statistical Offices: Pros and Cons for using Administrative records in Statistical Bureaus, *Fifty-fifth Plenary Session - Conference of European Statisticians*, 11-13 June 2007. Geneva.

United States Department of Agriculture (USDA). 2011. *Price Program: History, Methodology, Analysis, Estimates, and Dissemination*. USDA Publication: Washington, D.C.

United States Department of Health and Human Services (HHS). 2002. *Studies of Welfare Populations: Data Collection and Research Issues*. Publication of the U.S. Department of Health and Human Services: Washington, D.C.

2002. *Studies of Welfare Populations: Part II: Administrative Data*. Publication of the U.S. Department of Health and Human Services: Washington, D.C.

2002. *Studies of Welfare Populations: Access and Confidentiality Issues with Administrative Data*. Publication of the U.S. Department of Health and Human Services: Washington, D.C.

Wallgren, A. & Wallgren, B. 2007. *Register-based Statistics - Administrative Data for Statistical Purposes*. John Wiley & Sons: Chichester, UK.

2010. *Using administrative registers for agricultural statistics*. In Benedetti, B., Bee, M., Espa, G., & Piersimoni, F. (eds) *Agricultural Survey Methods*. John Wiley and Sons: Chichester, UK.

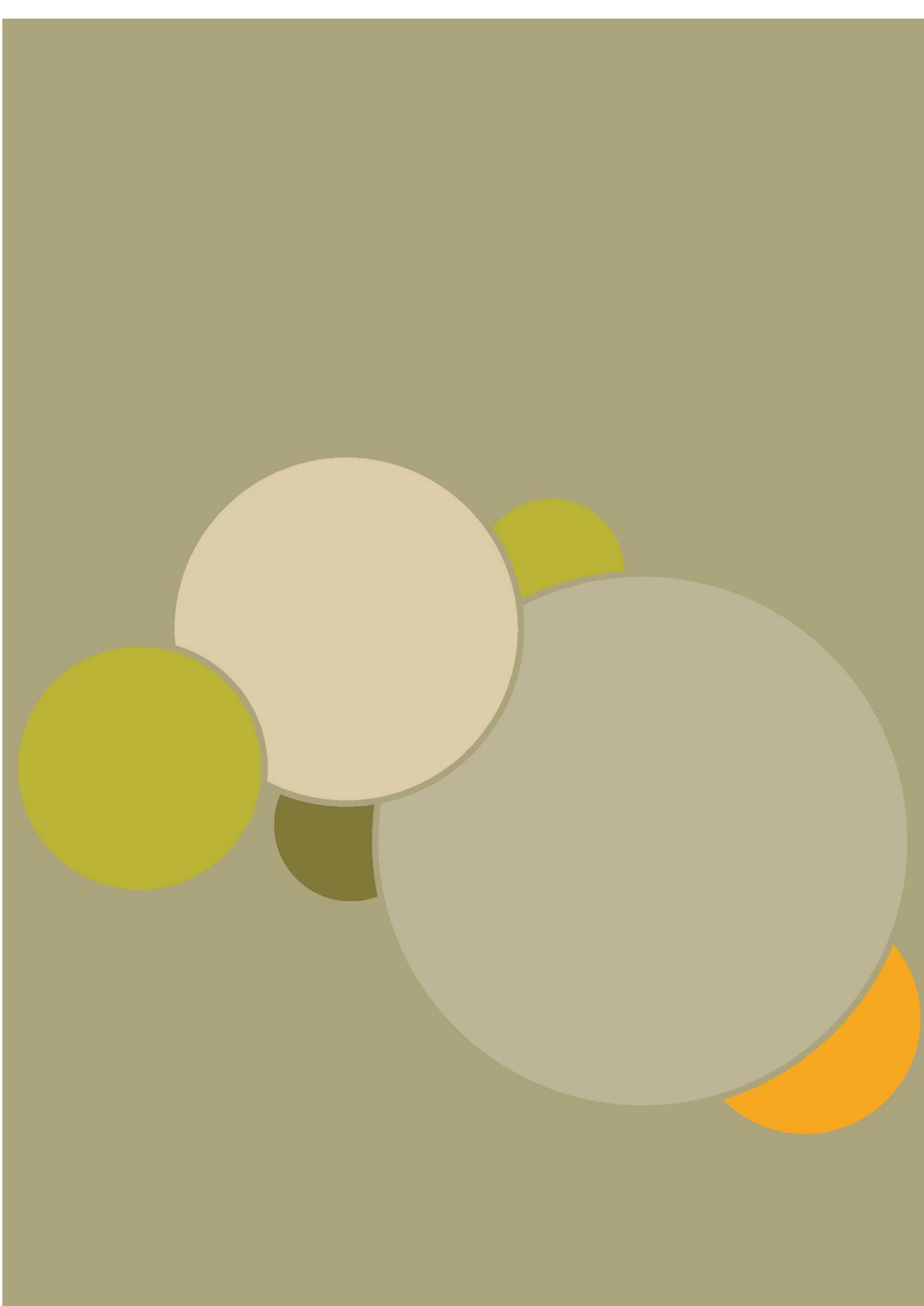
2011. *To understand the possibilities of administrative data, you must change your statistical paradigm*. In Proceedings of the American Statistical Association Section on Survey Research Methods, pp. 357365. American Statistical Association Publication: Alexandria, VA, USA.

Wang, J. & Fuller, W.A. 2003. The mean squared error of small area predictors constructed with estimated area variances. *Journal of the American Statistical Association*, 98(463): 716-723.

World Bank, FAO & UN. 2010. *Global Strategy to Improve Agricultural and Rural Statistics*. Report 56719-GL. World Bank Publication: Washington, D.C.

Yung, W., Rancourt, E. & Hidirolou, M. 2007. Administrative Data in Statistics Canada's business Surveys: The Present and the Future. Seminar on Registers in Statistics - methodology and quality, 21-23 May 2007, Helsinki.

Zhang, L. 2012. Topics of Statistical Theory for Register-based Statistics and Data Integration. *Statistica Neerlandica*, 66(1): 41-63.



ПРИЛОЖЕНИЯ

ПРИЛОЖЕНИЕ 1.

ЗАТРАТЫ НА РАЗРАБОТКУ АДМИНИСТРАТИВНЫХ ДАННЫХ

Предположения

1. В местных органах уровнями, которые имеют значение при составлении административных данных, являются районы, подрайоны и села/деревни.
 - a. Административные данные формируются на уровне села и консолидируются на уровне подрайонов для облегчения планирования на этом уровне.
 - b. Мобильные устройства для сбора данных используются на уровне села, и данные передаются на более высокие уровни.
2. Предполагается, что в стандартном районе имеется 10 подрайонов, а в каждом подрайоне - 50 деревень. Из каждой деревни выбираются 20 домашних хозяйств (10 домашних хозяйств для растениеводства и еще 10 домашних хозяйств для животноводства). Отсюда цифра $10 \times 50 \times 20 = 10\,000$ домашних хозяйств отбираются для сбора административных данных в округе.
3. На уровне сел и подрайонов назначаются работники по распространению сельскохозяйственных знаний, а на уровне районов работают три сотрудника (работник по распространению сельскохозяйственных знаний, работник по распространению знаний в области животноводства и статистик).
4. Требуется три мотоцикла на районном уровне и один - на уровне подрайона.
5. Период оценки системы составления административных данных составляет один год.
6. Невозможно было провести сравнение с бюджетом на обследование или перепись, поскольку этот бюджет представляет собой стоимость разработки административных данных при начале сбора административных данных. Очевидно, что затраты на разработку административных данных в последующие годы будут ниже, из-за исключения затрат на замену или обновление оборудования и обучение.

1 Пропаганда и коммуникация: для повышения информированности и осведомленности							
№	Описание	Комментарии	Количество	Частота: месяцы/дни/количества/число раз	Единица измерения	Стоимость единицы (шиллинги Уганды)	Всего
1	Разработка разъяснительных материалов	Консультации и подготовка материалов	1	30	Дни	194.03	5 820.90
2	Содействие чиновникам	Национальные и местные лидеры	20	15	Дни	104.48	31 343.28
3	Содействие по транспортировке	Транспорт	10	10	Автомобили	20.90	2 089.55
4	Привлечение местных общин		10	2	Лица	14.93	298.51
5	Объявления по радио	6 объявлений за три дня	60	3	Дни	11.94	2 149.25
6	Объявления о культовых местах	Один раз в 6-ти культовых местах	500	6	Дни	14.93	44 776.12
7	Ток-шоу	Один раз на двух радиостанциях	10	2	Дни	179.10	3 582.09
8	Брошюры и листовки	Консилидированно	1	1		2 985.07	2 985.07
9	Водитель	Водители	10	10	Дни	29.85	2 985.07
Подитог							96 029,85

2 Закупки							
2.1 Закупки транспортных средств							
№	Описание	Комментарии	Количество	Частота: месяцы/дни/количества/число раз	Единица измерения	Стоимость единицы (шиллинги Уганды)	Всего
1	Транспортные средства для района	3 мотоцикла	3	1	Время	2 388.06	7 164.18
2	Транспортные средства для подрайона	1 мотоцикл на каждый подрайон	10	1	Время	2 388.06	23 880.60
Подитог							31 044,78

2.2 Закупки оборудования для производства данных							
№	Описание	Комментарии	Количество	Частота: месяцы/дни/количества/число раз	Единица измерения	Стоимость единицы (шиллинги Уганды)	Всего
1	Закупка планшетов /смартфонов	Планшет закупается для каждого села	500	1	Разы	283.58	141 791.04
2	Закупка вопросников (твердые копии)	Каждому домашнему хозяйству в селе один вопросник на год	500	200	Домашние хозяйства	1.49	149 253.73
3	Закупка компьютеров с принтерами и сканерами	Каждый район и подрайон должен иметь один компьютер для управления данными	11	1	Компьютер	1 194.03	13 134.33
4	Закупка разного программного обеспечения	ПО для мобильных устройств и компьютеров	1	1	Единица	2 089.55	2 089.55
5	Закупка трафика в Интернете/данных	Планшеты и компьютеры	500	12	Месяцы	14.93	89 552.24
6	Закупка услуг программистов	Разработка и поддержка систем	10	2		74.63	1 492.54
Подитог							397 313.43

3 Обучение по формированию административных данных на разных уровнях							
№	Описание	Комментарии	Количество	Частота: месяцы/дни/количества/число раз	Единица измерения	Стоимость единицы (шиллинги Уганды)	Всего
1	Стипендия обучающимся	Обучение по 3 дня два раза в год	560	6	Дни	29.85	100 298.51
2	Учебные материалы		780	1	Лица	8.96	6 985.07
3	Оплата преподавателей	Обучение на уровне подрегионов	10	8	Преподаватели	358.21	28 656.72
4	Учебные площадки		10	6	Дни	149.25	8 955.22
5	Транспортные расходы		10	2	Разы	104.48	2 089.55
6	Расходы на связь		10	2	Разы	4.48	89.55
Подитог							147 074.63

4 Сбор/разработка административных данных на разных уровнях							
4.1 Заработная плата							
№	Описание	Комментарии	Количество	Частота: месяцы/дни/количества/число раз	Единица измерения	Стоимость единицы (шиллинги Уганды)	Всего
1	Разработка на уровне села	Зарплата специалистам по распространению сельскохозяйственных знаний в селах	500	12	Месяцы	208.96	1 253 731.34
2	Надзор на уровне подрайона	Зарплата специалистам по распространению сельскохозяйственных знаний в подрайоне	10	12	Месяцы	313.43	37 611.94
3	Надзор на уровне района	Зарплата специалистам по распространению сельскохозяйственных знаний в районе	3	12	Месяцы	417.91	15 044.78
Подитог							1 306 388.06

4.2 Транспорт в поле (сбор данных и мониторинг)							
№	Описание	Комментарии	Количество	Частота: месяцы/дни/количества/число раз	Единица измерения	Стоимость единицы (шиллинги Уганды)	Всего
1	Транспортные расходы на уровне подрайона	24 поездки	10	24	Поездки	29.85	7 164.18
2	Транспортные расходы на уровне района	12 поездок	10	12	Поездки	104.48	12 537.31
3	Визиты для мониторинга на национальном уровне	12 поездок	10	12	Поездки	417.91	50 149.25
Подитог							69 850.75

5 Анализ данных и отчетность							
№	Описание	Комментарии	Количество	Частота: месяцы/дни/количества/число раз	Единица измерения	Стоимость единицы (шиллинги Уганды)	Всего
1	Укрепление потенциала	Обучение, наставничество и привязанность	12	1	Обучение	597.01	7 164.18
2	Канцелярские товары		12	1	Канцелярские товары	149.25	1 791.04
3	Встречи по валидации данных	Совещание для обсуждения результатов с сотрудниками на уровне подрегиона	12	1		746.27	8 955.22
Подитог							17 910.45

6 Распространение данных участие в политике							
№	Описание	Комментарии	Количество	Частота: месяцы/дни/количества/число раз	Единица измерения	Стоимость единицы (шиллинги Уганды)	Всего
1	Производство и распространение отчетов	Распространение твердых копий, по Интернету и на семинарах	20	2	Разы	1 492.54	59 701.49
	Подитог						59 701.49
	ВСЕГО						2 125 313, 43

ПРИЛОЖЕНИЕ 2.

АСПЕКТЫ, АТТРИБУТЫ И ПОКАЗАТЕЛИ КАЧЕСТВА АДМИНИСТРАТИВНЫХ РЕГИСТРОВ

Аспекты качества	Атрибуты качества	Показатели качества
Источник административных данных	1. Актуальность	1.1. Полезность
		1.2. Использование по назначению
		1.3. Спрос на информацию
		1.4. Удовлетворенность основных пользователей
	2. Информационная безопасность и ограничения на использование информации	2.1. Законодательные рамки
		2.2. Защита персональных данных
		2.3. Ограничения в связи с регулированием конфиденциальности
		2.4. Соглашения о конфиденциальности
		2.5. Безопасная передача данных
		2.6. Конфиденциальность, целостность и наличие информации
		2.7. Защита данных
		2.8. Политика резервного копирования данных
	3. Обязательства по поставке данных	3.1. Затраты по поставке данных
		3.2. Соглашения о поставке
		3.3. Частота поставки
		3.4. Даты последних пяти поставок
		3.5. Пунктуальность
		3.6. Риски, связанные с отсутствием данных
		3.7. Разные методы для замены отсутствующих данных
		3.8. Средства доставки данных
		3.9. Формат файлов
		3.10. Отбор данных
	4. Контроль и постоянное улучшение	4.1. Сбор данных
		4.2. Контроль согласованности
		4.3. Контроль изменений
		4.4. Постоянное улучшение
	5. Обработка данных	5.1. Контроль единиц
		5.2. Контроль содержания переменных
		5.3. Контроль выбивающихся значений
5.4. Изменения		
5.5. Причины отсутствия изменений		
5.6. Изменения в связи с процедурой		
5.7. Использование систем управления данными		
5.8. Документация для баз данных		
5.9. Целостность баз данных		

Аспекты качества	Атрибуты качества	Показатели качества
Метаданные	1. Документация метаданных	1.1. Документация метаданных
	2. Полнота и ясность	2.1. Определение единиц совокупности
		2.2. Описание переменных
		2.3. Сообщение об изменениях в определениях/концепциях
3. Использование уникальных ключей	3.1. Ключи идентификации	
	3.2. Сопоставимость и идентификация ключей	
	3.3. Уникальная комбинация переменных	
4. Сопоставимость	4.1. Сопоставимость определения единиц	
	4.2. Сопоставимость определения переменных	
Данные	1. Технический контроль	1.1. Читаемые данные
		1.2. Переопределение концепций и метаданных в случае более одного источника
		1.3. Соответствие между данными и метаданными
		1.4. Метод увязки записей
		1.5. Проверка эффективности метода увязки записей
	2. Охват	2.1. Избыточный охват
		2.2. Ошибки классификации
	3. Увязка записей	3.1. Процент увязки записей
	4. Полнота	4.1. Процент неответа единиц наблюдения
		4.2. Процент неответа на вопросы обследования
	5. Измерение	5.1. Внешний контроль (аудит)
	6. Ключи идентификации	6.1. Доля записей с уникальным кодом
	7. Обработка данных	7.1. Редактирование данных
		7.2. Импутация
	8. Точность данных	8.1 Точность данных
	9. Кодирование	9.1. Использование стандартного кодирования
		9.2. Проверка кодирования
		9.3. Процент ошибок кодирования
		9.4. Процент записей без кодов
	10. Свежесть данных	10.1. Более 90% единиц, созданных в течение года t, были зарегистрированы до конца года t+1
11. Повторные записи	11.1. Процент повторных записей одной и той же единицы	
12. Другие контроли	12.1. Количество единиц с действительными значениями в идентификационных ключах	
	12.2. Таблицы статистической операции были проверены с помощью автоматических процедур	

Аспекты качества	Атрибуты качества	Показатели качества
Статистический продукт	1. Сопоставимость	1.1. Длина сопоставимых временных рядов
		1.2. Сопоставимость микроданных во времени
	2. Актуальность	2.1. Идентификация пользователей
		2.2. Информация о пользователях
		2.3. Степень удовлетворенности конечного пользователя
		2.4. Полезность (использование по назначению)
	3. Обоснованность	3.1. Согласованность статистики с разной периодичностью
		3.2. Согласованность статистики одинакового социально-экономического масштаба
	4. Доступность и ясность	4.1. Доступность по Интернету
		4.2. Степень полноты метаданных
	5. Точность	5.1. Коэффициент вариации
		5.2. Процент ответов единиц
		5.3. Процент ответов на вопросы обследования
		5.4. Процент импутированных данных
		5.5. Процент редактирования
		5.6. Процент избыточного охвата
		5.7. Процент ошибок классификации
	6. Своевременность и пунктуальность	6.1. Пунктуальность распространения статистического продукта
		6.2. Период времени между событием или явлением и доступностью информации
		6.3. Период времени между временем учета в административных данных и даты доступности данных для статистического управления

ПРИЛОЖЕНИЕ 3

ОБЗОР ПРОБЕЛОВ В АДМИНИСТРАТИВНЫХ ДАННЫХ ПО СЕЛЬСКОМУ ХОЗЯЙСТВУ И ПУТИ УЛУЧШЕНИЯ.

Раздел/Тема	Пробел/слабость	Возможные решения
Институциональ- ные рамки	Плохая координация между НСС и различными агентствами, занимающимися сбором административных данных по сельскому хозяйству и управлением ими.	Индия дает хороший пример того, как координировать различные федеральные и государственные учреждения.
	Слабая актуальность данных из-за отсутствия связи между производителями данных и пользователями.	Форумы для общения с заинтересованными сторонами, такими как Национальный технический комитет по сельскохозяйственной статистике. Укрепление комитетов производителей и пользователей и проведение регулярных семинаров по распространению данных.
	Недостаточная физическая и статистическая инфраструктура.	Техническая поддержка для улучшения институциональной инфраструктуры. Хороший опыт Индии, где имеются подробные кадастровые карты, часто обновляемые земельные записи и институт постоянных сельских агентства по отчетности.
	Отсутствие обучения или плохое обучение и надзор. Показатели качества отражают плохое обучение.	Предоставьте надлежащие руководящие принципы, например, иерархические системы обучения, а также передовой опыт по надзору и поддержке для ARDS Объединенной Республики Танзания. Также обратитесь к специальному обучению и строгим процедурам контроля качества в Индии.
Ресурсы	Во многих министерствах, департаментах и агентствах в сельскохозяйственном секторе отсутствуют статистические отделы или подразделения.	Создание статистических единиц.
	Недостаточное количество квалифицированного персонала и низкий уровень удержания персонала в основном из-за низких зарплат и плохих условий труда.	Оценка потребностей в людских ресурсах и включение средств, необходимых для набора и обучения персонала, в национальный бюджет и бюджет министерства, полугосударственных учреждений и местных органов власти.
	Плохие структуры стимулирования среди работников, ведущие к плохому поведению, например, уклонение от работы среди работников учреждений.	Разработайте эффективный план вознаграждения и стимулирования для мотивации сотрудников на основе их фактической эффективности.
	Сотрудники по распространению сельскохозяйственных знаний и/или руководители, которые часто собирают административные сельскохозяйственные данные, также выполняют множество других функций.	Обозначьте потребности в бюджете и объединяйте ресурсы (уменьшите пересечения, рационализируйте действия, чтобы ясно разъяснить сбор данных в описании работы), для сокращения расходов.

Раздел/Тема	Пробел/слабость	Возможные решения
Финансовые ресурсы, источники финансирования и стратегии устойчивости	Недостаточные и неустойчивые финансовые ресурсы частично из-за выделения очень небольших средств на сельское хозяйство в национальных бюджетах.	Увеличение бюджетных ассигнований в национальных бюджетах и уделение внимания административным сельскохозяйственным данным.
	Большая часть финансирования мероприятий, как правило, осуществляется донорскими агентствами, имеющими короткий срок жизни. После прекращения донорского финансирования системы обычно рушатся.	То же, что и выше, а также пропаганда передового опыта - лоббирование правительства в Азиатско-Тихоокеанском регионе было успешно для получения доступа к национальному бюджету для систем регулярного сбора данных.
	Нерациональное использование ресурсов, учитывая дублирование деятельности по сбору данных различными учреждениями.	Лучшая координация между министерствами, департаментами и агентствами, объединение ресурсов (человеческих и финансовых) и согласование работы с административными данными позволит сократить расходы.
	Отсутствие информации по эффективности затрат в системе сбора сельскохозяйственных данных.	Определить затраты на сбор данных, анализ, распространение, управление базой данных.
Сбор данных и управление данными	Плохие инструменты сбора данных – вопросники и руководства.	Передовой опыт Индии и Объединенной Республики Танзании в отношении ясных вопросников и пособий с четкими указаниями о том, кто / как / когда должен заполнять различные формы.
	Субъективная отчетность о посевных площадях, а также фактических и прогнозных урожаях.	Измерение площади и оценка производства. Разработать системы регистрации земли, такие как в Индии, и программы, которые побуждают фермеров вести учет.
	Оценка производства продовольственных и второстепенных культур.	Ищите новые потенциальные источники данных и специальные исследования и / или опросы.
	Пробелы в методах оценивания при смешанном и непрерывном выращивании сельскохозяйственных культур.	Ожидайте уроков из исследования Глобальной стратегии « <i>Методологии оценки посевных площадей и урожайности при смешанном и непрерывном выращивании с/х культур</i> », проведенного Индийским научно-исследовательским институтом сельскохозяйственной статистики (IASRI).
Статистическое программное обеспечение и ИКТ	Применение статистических стандартов и методов и использование современных технологий.	Внедрить и расширить использование современных технологий, таких как оборудование GPS, мобильный телефон, КПК, сканеры, дистанционное зондирование и т. д., адаптированных для процессов обработки административных данных в сельском хозяйстве. Обучите и вооружите персонал, чтобы использовать эти технологии для сбора, обработки и анализа данных, их компиляции, хранения и распространения.
Наличие определенных данных	Часто неполные, устаревшие, противоречивые и ненадежные данные по сельскохозяйственным животным и использование неоднородных форматов в местных администрациях.	Улучшить навыки обработки данных, нехватки ресурсов и т.д. Гармонизация форматов при сборе и анализе данных. В Объединенной Республике Танзания, Уганде и Мозамбике существует ряд передовых методов.
	Данные по другим основным элементам данных, таким как лесное хозяйство, рыболовство, сельскохозяйственные ресурсы, также являются неполными.	Ищите новые потенциальные источники, например, фермерские ассоциации и производители сырья.
	Сельскохозяйственные культуры: второстепенные культуры и системы непрерывного земледелия создают проблемы.	Специальные исследования / опросы по мелким культурам и результаты исследования IASRI по сплошному посеву..
	Торговые данные / экспорт и импорт: Таможня предоставляет официальную статистику торговли, но неформальная торговля не зарегистрирована.	Проводите обследования неформальной трансграничной торговли.
	Земной покров: определения и классификации неадекватны.	Улучшите координацию для определения релевантных категорий.
	Наличие данных: доступ затруднен.	Используйте CountrySTAT.

Раздел/Тема	Пробел/слабость	Возможные решения
Качество данных	Расхождение между различными показателями одного и того же элемента - отсутствие последовательности, согласованности и сопоставимости.	Повышение конкретности определений, заимствование идей для обучения и контроля из Индии, мониторинг качества административных данных, как в Индии и Мексике. Схема улучшения статистики сельскохозяйственных культур (ICS) предусматривает надзор за сборщиками данных для проверки точности исходных данных. Схема своевременной отчетности (TRS) - это попытка улучшить своевременность данных.
Использование административных данных, особенно в развитых странах	Многие виды использования не применяются широко в развивающихся странах.	Адаптировать различные виды использования к развивающимся странам и применять их.
	Объединение данных из многих источников. Несогласованные данные публикуются без всяких разъяснений. Увязка записей не используется.	Методология увязки записей и оценки ошибки измерения необходима для поддержания высококачественных баз данных, которые объединяют много файлов административных данных. Объединение данных переписи и административных данных для создания эффективных совокупностей обследования, статистических моделей и процессов ручного анализа и вероятностной увязки записей. Программное обеспечение также имеется. Есть также примеры из Индии, Мозамбика и Уганды.
	Оценка малых областей.	Используйте административные данные как ковариаты при построении оценок малых областей и прогнозов на основе моделей. Уроки из опыта Эфиопии.
	Повышение эффективности оценок на базе обследования, то есть предоставляйте контрольные итоги по совокупности.	Используйте калибровку (Carfagna and Carfagna, 2010).
	Улучшение построения выборки и дизайна выборочного обследования.	Стройте территориальные совокупности для выборки.

ПРИЛОЖЕНИЕ 4.

ОБРАЗЕЦ МЕМОРАНДУМА О ВЗАИМОПОНИМАНИИ МЕЖДУ НАЦИОНАЛЬНОЙ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИЧЕСКОЙ СЛУЖБОЙ И НАЦИОНАЛЬНОЙ СЛУЖБОЙ СОХРАНЕНИЯ РЕСУРСОВ США

**МЕМОРАНДУМ О ВЗАИМОПОНИМАНИИ
МИНИСТЕРСТВА СЕЛЬСКОГО ХОЗЯЙСТВА США
между
НАЦИОНАЛЬНОЙ СЛУЖБОЙ СОХРАНЕНИЯ ПРИРОДНЫХ РЕСУРСОВ
и
НАЦИОНАЛЬНОЙ СЕЛЬСКОХОЗЯЙСТВЕННОЙ СТАТИСТИЧЕСКОЙ СЛУЖБОЙ
ОТНОСИТЕЛЬНО ОБМЕНА ДАННЫМИ И СТАТИСТИЧЕСКОЙ ИНФОРМАЦИЕЙ
И СТРАТЕГИЙ РАСПРОСТРАНЕНИЯ**

ЮРИДИЧЕСКАЯ ОСНОВА

Закон о сельскохозяйственном маркетинге 1946 года, заголовки Свода законов США 7,1621-6127 и 2204g, Закон о регулировании сельского хозяйства 1938 года (Закон 1938 года), Закон о сохранении почвы 1935 года с поправками и Закон 1985 года о продовольственной безопасности с поправками.

ЦЕЛЬ

Настоящий Меморандум о взаимопонимании (МОВ) устанавливает основу для взаимоотношений между Службой сохранения природных ресурсов и Национальной службой сельскохозяйственной статистики (NASS) Министерства сельского хозяйства США (USDA) в поддержку деятельности Министерства сельского хозяйства США. Этот меморандум о взаимопонимании помогает развивать и совершенствовать обмен данными и информацией о фермах и сельском хозяйстве страны, чтобы служить интересам Министерства сельского хозяйства США, сельскохозяйственного сообщества и нации в целом.

Оба агентства хотят, чтобы действия в соответствии с этим Меморандумом о взаимопонимании помогли бы стимулировать и расширять обмен стратегиями (как традиционными, так и нетрадиционными) для охвата и обслуживания недостаточно обслуживаемых сельскохозяйственных сообществ в стране, включая американских индейцев, коренных жителей Аляски и индейских племен. Этот меморандум о взаимопонимании поможет расширить доступ в эти сообщества, чтобы обеспечить доверие и присутствие NRCS и NASS, а также расширить участие в программах и услугах NRCS и NASS.

ИСТОРИЯ ВОПРОСА

NRCS и NASS имеют долгую историю сотрудничества и координации. В обоих агентствах есть организации, которые предоставляют свой основной канал коммуникации общественности и предлагают данные и информацию национальным операторам фермерских хозяйств и сельскохозяйственному бизнесу.

Как NRCS, так и NASS занимаются экологическими вопросами в отношении природных ресурсов страны. NRCS предоставляет техническую экспертизу и помощь. NASS и NRCS суммируют те положительные шаги, которые предпринимаются в США, и сообщают о них для демонстрации экологически ответственных решений в сельскохозяйственных операциях.

NRCS зависит от NASS в части предоставления оценок урожайности сельскохозяйственных культур по округам, штатам и по стране в целом для планирования и реализации мероприятий. NASS зависит от NRCS в части предоставления данных ферм о сохранении сельскохозяйственных земель и другой информации о программах, используемой в качестве вторичных источников данных для получения оценок NASS и для обновления списка операторов ферм NASS данными, используемыми для целей выборки. NASS и NRCS имеют опыт сотрудничества в таких проектах, как влияние и оценка сохранения земель, а также национальная инвентаризация ресурсов.

ОТВЕТСТВЕННОСТЬ

В соответствии с данным документом, оба агентства поделятся стратегиями распространения, чтобы коллективно интегрировать клиентов USDA, обеспечить лучшие коммуникации, расширить доступность и стать лучшими ресурсами для всех клиентов USDA.

Доступ и использование данных NRCS и NASS будут соответствовать закону или постановлению, разрешающему сбор данных, а также будут соответствовать руководящим принципам и стандартам качества данных Административно-бюджетного управления.

NRCS:

- Будет помогать NASS в разработке и ведении текущих списков операторов ферм, предоставляя рекомендации по использованию файлов базы данных на уровнях производителей;
- Разрешит NASS электронный доступ и хранение данных NRCS, к которым NASS предоставлен прямой доступ;
- Предоставит информацию на уровне производителя для NASS и счетчиков Национальной ассоциации департаментов сельского хозяйства штатов (NASDA), с которыми NASS заключила контракт, из различных файлов базы данных NRCS, содержащих имена, адреса и связанную информацию, для использования NASS в текущих программах статистики;
- Включит описание усилий NASS в области обновления сельскохозяйственной переписи в брифинги и презентации на местном уровне, уровне штатов и национальном уровне;
- Окажет помощь в информировании всех клиентов о сельскохозяйственной переписи;
- Используя связи между племенами, где были установлены доверительные отношения с индейскими и другими племенами, познакомит сотрудников NASS и счетчиков NASDA с племенными образованиями и общественными организациями и объяснит важность сбора точных данных сельскохозяйственной переписи;
- Сохранит конфиденциальность данных NASS.

NASS:

- Предоставит NRCS текущие и исторические статистические данные, опубликованные и по запросу;
- Сохранит конфиденциальность данных NRCS, к которым осуществляется доступ;
- Предоставит информационные рассылки по себестоимости целевым группам фермеров для уведомления о новых или модифицированных программах NRCS;

- Будет работать с NRCS во время разработки содержания сельскохозяйственной переписи, чтобы определить потребности в данных NRCS для распределения средств программы на основе статистики и сегментного анализа рынка;
- Предпримет совместные усилия, чтобы помочь NRCS увеличить помощь социально незащищенным и недостаточно обслуживаемым фермерам и скотоводам;
- Предоставит помощь NRCS в отношении статистической интерпретации данных, полученных и разработанных в ходе обследований NASS и проведения переписи;
- Представит деятельность NRCS в брифингах и презентациях на местном, региональном и национальном уровнях для укрепления партнерских отношений и услуг USDA.

ВЗАИМОПОНИМАНИЕ

NRCS, NASS и их соответствующие офисы будут заниматься своими видами деятельности и использовать свои собственные ресурсы, в том числе расходование собственных средств, для достижения этих целей. Каждая сторона будет осуществлять свою деятельность в согласованной и взаимовыгодной форме.

Ничто в этом МОВ не обязывает NRCS или NASS выделять или переводить какие-либо средства. Конкретные рабочие проекты или виды деятельности, которые подразумевают передачу средств, услуг или имущества между учреждениями и офисами, потребуют выполнения отдельных соглашений и будут зависеть от наличия соответствующих средств. Такая деятельность должна быть санкционирована соответствующим правомочным органом. Настоящий МОВ не предоставляет таких полномочий.

Настоящий Меморандум вступает в силу с момента подписания представителями агентства USDA и действует в течение не более 5 лет с даты заключения. Этот Меморандум может быть расширен или изменен по письменному запросу или последующему письменному согласию другой стороны. NRCS или NASS могут прекратить действие настоящего Меморандума с 60-дневным письменным уведомлением другой стороны.

NASS и NRCS будут координировать, по крайней мере, ежегодно, какие конкретные данные и информация будут передаваться в соответствии с целью настоящего Меморандума.

NASS и NRCS будут координировать сотрудничество между офисами NRCS в штатах и полевыми офисами NASS для получения доступа к необходимым данным и информации для достижения целей настоящего Меморандума. Копия этого Соглашения будет распространена в офисы директоров NASS в штатах и офисы директоров NRCS в штатах.

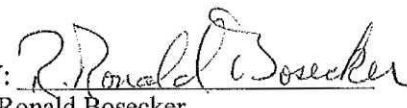
IT IS SO AGREED

FOR NATURAL RESOURCES
CONSERVATION SERVICE

BY: 
Arlen L. Lancaster
Chief, Natural Resources
Conservation Service

DATE: APR - 6 2007

FOR NATIONAL
AGRICULTURAL
STATISTICS SERVICE

BY: 
R. Ronald Bosecker
Administrator, National Agricultural
Statistics Service

DATE: 4/17/07

