## Contents

<table>
<thead>
<tr>
<th>List of Contributors</th>
<th>xvii</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introduction</strong></td>
<td>xxi</td>
</tr>
</tbody>
</table>

1. The present state of agricultural statistics in developed countries: situation and challenges  
1.1 Introduction 1  
1.2 Current state and political and methodological context 4  
  1.2.1 General 4  
  1.2.2 Specific agricultural statistics in the UNECE region 6  
1.3 Governance and horizontal issues 15  
  1.3.1 The governance of agricultural statistics 15  
  1.3.2 Horizontal issues in the methodology of agricultural statistics 16  
1.4 Development in the demand for agricultural statistics 20  
1.5 Conclusions 22  
Acknowledgements 23  
Reference 24

**Part I  Census, Frames, Registers and Administrative Data  25**

2. Using administrative registers for agricultural statistics 27  
  2.1 Introduction 27  
  2.2 Registers, register systems and methodological issues 28  
  2.3 Using registers for agricultural statistics 29  
    2.3.1 One source 29  
    2.3.2 Use in a farm register system 30  
    2.3.3 Use in a system for agricultural statistics linked with the business register 30  
  2.4 Creating a farm register: the population 34  
  2.5 Creating a farm register: the statistical units 38  
  2.6 Creating a farm register: the variables 42
# CONTENTS

## 3 Alternative sampling frames and administrative data. What is the best data source for agricultural statistics?

3.1 Introduction 45  
3.2 Administrative data 46  
3.3 Administrative data versus sample surveys 46  
3.4 Direct tabulation of administrative data 46  
3.4.1 Disadvantages of direct tabulation of administrative data 47  
3.5 Errors in administrative registers 48  
3.5.1 Coverage of administrative registers 48  
3.6 Errors in administrative data 49  
3.6.1 Quality control of the IACS data 49  
3.6.2 An estimate of errors of commission and omission in the IACS data 50  
3.7 Alternatives to direct tabulation 51  
3.7.1 Matching different registers 51  
3.7.2 Integrating surveys and administrative data 52  
3.7.3 Taking advantage of administrative data for censuses 52  
3.7.4 Updating area or point sampling frames with administrative data 53  
3.8 Calibration and small-area estimators 53  
3.9 Combined use of different frames 54  
3.9.1 Estimation of a total 55  
3.9.2 Accuracy of estimates 55  
3.9.3 Complex sample designs 56  
3.10 Area frames 57  
3.10.1 Combining a list and an area frame 57  
3.11 Conclusions 58  
Acknowledgements 59  
References 60

## 4 Statistical aspects of a census

4.1 Introduction 63  
4.2 Frame 64  
4.2.1 Coverage 64  
4.2.2 Classification 64  
4.2.3 Duplication 65  
4.3 Sampling 65  
4.4 Non-sampling error 66  
4.4.1 Response error 66  
4.4.2 Non-response 67  
4.5 Post-collection processing 68  
4.6 Weighting 68  
4.7 Modelling 69  
4.8 Disclosure avoidance 69  
4.9 Dissemination 70
5 Using administrative data for census coverage 73
5.1 Introduction 73
5.2 Statistics Canada’s agriculture statistics programme 74
5.3 1996 Census 75
5.4 Strategy to add farms to the farm register 75
  5.4.1 Step 1: Match data from E to M 76
  5.4.2 Step 2: Identify potential farm operations among the unmatched records from E 76
  5.4.3 Step 3: Search for the potential farms from E on M 76
  5.4.4 Step 4: Collect information on the potential farms 77
  5.4.5 Step 5: Search for the potential farms with the updated key identifiers 77
5.5 2001 Census 77
  5.5.1 2001 Farm Coverage Follow-up 77
  5.5.2 2001 Coverage Evaluation Study 77
5.6 2006 Census 78
  5.6.1 2006 Missing Farms Follow-up 79
  5.6.2 2006 Coverage Evaluation Study 80
5.7 Towards the 2011 Census 81
5.8 Conclusions 81
Acknowledgements 83
References 83

Part II Sample Design, Weighting and Estimation 85
6 Area sampling for small-scale economic units 87
  6.1 Introduction 87
  6.2 Similarities and differences from household survey design 88
    6.2.1 Probability proportional to size selection of area units 88
    6.2.2 Heterogeneity 90
    6.2.3 Uneven distribution 90
    6.2.4 Integrated versus separate sectoral surveys 90
    6.2.5 Sampling different types of units in an integrated design 91
  6.3 Description of the basic design 91
  6.4 Evaluation criterion: the effect of weights on sampling precision 93
    6.4.1 The effect of ‘random’ weights 93
    6.4.2 Computation of $D^2$ from the frame 94
    6.4.3 Meeting sample size requirements 94
  6.5 Constructing and using ‘strata of concentration’ 95
    6.5.1 Concept and notation 95
    6.5.2 Data by StrCon and sector (aggregated over areas) 95
    6.5.3 Using StrCon for determining the sampling rates: a basic model 97
  6.6 Numerical illustrations and more flexible models 97
6.6.1 Numerical illustrations 97
6.6.2 More flexible models: an empirical approach 100
6.7 Conclusions 104
Acknowledgements 105
References 105

7 On the use of auxiliary variables in agricultural survey design 107
7.1 Introduction 107
7.2 Stratification 109
7.3 Probability proportional to size sampling 113
7.4 Balanced sampling 116
7.5 Calibration weighting 118
7.6 Combining ex ante and ex post auxiliary information: a simulated approach 124
7.7 Conclusions 128
References 129

8 Estimation with inadequate frames 133
8.1 Introduction 133
8.2 Estimation procedure 133
8.2.1 Network sampling 133
8.2.2 Adaptive sampling 135
References 138

9 Small-area estimation with applications to agriculture 139
9.1 Introduction 139
9.2 Design issues 140
9.3 Synthetic and composite estimates 140
9.3.1 Synthetic estimates 141
9.3.2 Composite estimates 141
9.4 Area-level models 142
9.5 Unit-level models 144
9.6 Conclusions 146
References 147

Part III GIS and Remote Sensing 149

10 The European land use and cover area-frame statistical survey 151
10.1 Introduction 151
10.2 Integrating agricultural and environmental information with LUCAS 154
10.3 LUCAS 2001–2003: Target region, sample design and results 155
10.4 The transect survey in LUCAS 2001–2003 156
10.5 LUCAS 2006: a two-phase sampling plan of unclustered points 158
10.6 Stratified systematic sampling with a common pattern of replicates 159
10.7 Ground work and check survey 159
10.8 Variance estimation and some results in LUCAS 2006 160
10.9 Relative efficiency of the LUCAS 2006 sampling plan 161
10.10 Expected accuracy of area estimates with the LUCAS 2006 scheme 163
10.11 Non-sampling errors in LUCAS 2006 164
  10.11.1 Identification errors 164
  10.11.2 Excluded areas 164
10.12 Conclusions 165
Acknowledgements 166
References 166

11 Area frame design for agricultural surveys 169
11.1 Introduction 169
  11.1.1 Brief history 170
  11.1.2 Advantages of using an area frame 171
  11.1.3 Disadvantages of using an area frame 171
  11.1.4 How the NASS uses an area frame 172
11.2 Pre-construction analysis 173
11.3 Land-use stratification 176
11.4 Sub-stratification 178
11.5 Replicated sampling 180
11.6 Sample allocation 183
11.7 Selection probabilities 185
  11.7.1 Equal probability of selection 186
  11.7.2 Unequal probability of selection 187
11.8 Sample selection 188
  11.8.1 Equal probability of selection 188
  11.8.2 Unequal probability of selection 188
11.9 Sample rotation 189
11.10 Sample estimation 190
11.11 Conclusions 192

12 Accuracy, objectivity and efficiency of remote sensing for agricultural statistics 193
12.1 Introduction 193
12.2 Satellites and sensors 194
12.3 Accuracy, objectivity and cost-efficiency 195
12.4 Main approaches to using EO for crop area estimation 196
12.5 Bias and subjectivity in pixel counting 197
12.6 Simple correction of bias with a confusion matrix 197
12.7 Calibration and regression estimators 197
12.8 Examples of crop area estimation with remote sensing in large regions 199
  12.8.1 US Department of Agriculture 199
  12.8.2 Monitoring agriculture with remote sensing 200
  12.8.3 India 200
12.9 The GEOSS best practices document on EO for crop area estimation 200
12.10 Sub-pixel analysis 201
12.11 Accuracy assessment of classified images and land cover maps 201
12.12 General data and methods for yield estimation 203
12.13 Forecasting yields 203
12.14 Satellite images and vegetation indices for yield monitoring 204
12.15 Examples of crop yield estimation/forecasting with remote sensing 205
  12.15.1 USDA 205
  12.15.2 Global Information and Early Warning System 206
  12.15.3 Kansas Applied Remote Sensing 207
  12.15.4 MARS crop yield forecasting system 207

References 207

13 Estimation of land cover parameters when some covariates are missing 213
  13.1 Introduction 213
  13.2 The AGRIT survey 214
    13.2.1 Sampling strategy 214
    13.2.2 Ground and remote sensing data for land cover estimation in a small area 216
  13.3 Imputation of the missing auxiliary variables 218
    13.3.1 An overview of the missing data problem 218
    13.3.2 Multiple imputation 219
    13.3.3 Multiple imputation for missing data in satellite images 221
  13.4 Analysis of the 2006 AGRIT data 222
  13.5 Conclusions 227

References 229

Part IV Data Editing and Quality Assurance 231

14 A generalized edit and analysis system for agricultural data 233
  14.1 Introduction 233
  14.2 System development 236
    14.2.1 Data capture 236
    14.2.2 Edit 237
    14.2.3 Imputation 238
  14.3 Analysis 239
    14.3.1 General description 239
    14.3.2 Micro-analysis 239
    14.3.3 Macro-analysis 240
  14.4 Development status 240
  14.5 Conclusions 241

References 242

15 Statistical data editing for agricultural surveys 243
  15.1 Introduction 243
  15.2 Edit rules 245
  15.3 The role of automatic editing in the editing process 246
  15.4 Selective editing 247
    15.4.1 Score functions for totals 248
    15.4.2 Score functions for changes 250
    15.4.3 Combining local scores 251
    15.4.4 Determining a threshold value 252
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.5</td>
<td>An overview of automatic editing</td>
<td>253</td>
</tr>
<tr>
<td>15.6</td>
<td>Automatic editing of systematic errors</td>
<td>255</td>
</tr>
<tr>
<td>15.7</td>
<td>The Fellegi–Holt paradigm</td>
<td>256</td>
</tr>
<tr>
<td>15.8</td>
<td>Algorithms for automatic localization of random errors</td>
<td>257</td>
</tr>
<tr>
<td>15.8.1</td>
<td>The Fellegi–Holt method</td>
<td>257</td>
</tr>
<tr>
<td>15.8.2</td>
<td>Using standard solvers for integer programming problems</td>
<td>259</td>
</tr>
<tr>
<td>15.8.3</td>
<td>The vertex generation approach</td>
<td>259</td>
</tr>
<tr>
<td>15.8.4</td>
<td>A branch-and-bound algorithm</td>
<td>260</td>
</tr>
<tr>
<td>15.9</td>
<td>Conclusions</td>
<td>263</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>264</td>
</tr>
<tr>
<td>16</td>
<td>Quality in agricultural statistics</td>
<td>267</td>
</tr>
<tr>
<td>16.1</td>
<td>Introduction</td>
<td>267</td>
</tr>
<tr>
<td>16.2</td>
<td>Changing concepts of quality</td>
<td>268</td>
</tr>
<tr>
<td>16.2.1</td>
<td>The American example</td>
<td>268</td>
</tr>
<tr>
<td>16.2.2</td>
<td>The Swedish example</td>
<td>271</td>
</tr>
<tr>
<td>16.3</td>
<td>Assuring quality</td>
<td>274</td>
</tr>
<tr>
<td>16.3.1</td>
<td>Quality assurance as an agency undertaking</td>
<td>274</td>
</tr>
<tr>
<td>16.3.2</td>
<td>Examples of quality assurance efforts</td>
<td>275</td>
</tr>
<tr>
<td>16.4</td>
<td>Conclusions</td>
<td>276</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>276</td>
</tr>
<tr>
<td>17</td>
<td>Statistics Canada’s Quality Assurance Framework applied to agricultural statistics</td>
<td>277</td>
</tr>
<tr>
<td>17.1</td>
<td>Introduction</td>
<td>277</td>
</tr>
<tr>
<td>17.2</td>
<td>Evolution of agriculture industry structure and user needs</td>
<td>278</td>
</tr>
<tr>
<td>17.3</td>
<td>Agriculture statistics: a centralized approach</td>
<td>279</td>
</tr>
<tr>
<td>17.4</td>
<td>Quality Assurance Framework</td>
<td>281</td>
</tr>
<tr>
<td>17.5</td>
<td>Managing quality</td>
<td>283</td>
</tr>
<tr>
<td>17.5.1</td>
<td>Managing relevance</td>
<td>283</td>
</tr>
<tr>
<td>17.5.2</td>
<td>Managing accuracy</td>
<td>286</td>
</tr>
<tr>
<td>17.5.3</td>
<td>Managing timeliness</td>
<td>293</td>
</tr>
<tr>
<td>17.5.4</td>
<td>Managing accessibility</td>
<td>294</td>
</tr>
<tr>
<td>17.5.5</td>
<td>Managing interpretability</td>
<td>296</td>
</tr>
<tr>
<td>17.5.6</td>
<td>Managing coherence</td>
<td>297</td>
</tr>
<tr>
<td>17.6</td>
<td>Quality management assessment</td>
<td>299</td>
</tr>
<tr>
<td>17.7</td>
<td>Conclusions</td>
<td>300</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td></td>
<td>300</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>300</td>
</tr>
</tbody>
</table>

**Part V**

Data Dissemination and Survey Data Analysis 303

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>The data warehouse: a modern system for managing data</td>
<td>305</td>
</tr>
<tr>
<td>18.1</td>
<td>Introduction</td>
<td>305</td>
</tr>
<tr>
<td>18.2</td>
<td>The data situation in the NASS</td>
<td>306</td>
</tr>
<tr>
<td>18.3</td>
<td>What is a data warehouse?</td>
<td>308</td>
</tr>
<tr>
<td>18.4</td>
<td>How does it work?</td>
<td>308</td>
</tr>
</tbody>
</table>
18.5 What we learned 310
18.6 What is in store for the future? 312
18.7 Conclusions 312

19 Data access and dissemination: some experiments during the First National Agricultural Census in China 313
19.1 Introduction 313
19.2 Data access and dissemination 314
19.3 General characteristics of SDA 316
19.4 A sample session using SDA 318
19.5 Conclusions 320
References 322

20 Analysis of economic data collected in farm surveys 323
20.1 Introduction 323
20.2 Requirements of sample surveys for economic analysis 325
20.3 Typical contents of a farm economic survey 326
20.4 Issues in statistical analysis of farm survey data 327
  20.4.1 Multipurpose sample weighting 327
  20.4.2 Use of sample weights in modelling 328
20.5 Issues in economic modelling using farm survey data 330
  20.5.1 Data and modelling issues 330
  20.5.2 Economic and econometric specification 331
20.6 Case studies 332
  20.6.1 ABARE broadacre survey data 332
  20.6.2 Time series model of the growth in fodder use in the Australian cattle industry 333
  20.6.3 Cross-sectional model of land values in central New South Wales 335
References 338

21 Measuring household resilience to food insecurity: application to Palestinian households 341
21.1 Introduction 341
21.2 The concept of resilience and its relation to household food security 343
  21.2.1 Resilience 343
  21.2.2 Households as (sub) systems of a broader food system, and household resilience 345
  21.2.3 Vulnerability versus resilience 345
21.3 From concept to measurement 347
  21.3.1 The resilience framework 347
  21.3.2 Methodological approaches 348
21.4 Empirical strategy 350
  21.4.1 The Palestinian data set 350
  21.4.2 The estimation procedure 351
21.5 Testing resilience measurement 359
  21.5.1 Model validation with CART 359
  21.5.2 The role of resilience in measuring vulnerability 363
21.5.3 Forecasting resilience 364
21.6 Conclusions 365
References 366

22 Spatial prediction of agricultural crop yield 369
22.1 Introduction 369
22.2 The proposed approach 372
  22.2.1 A simulated exercise 374
22.3 Case study: the province of Foggia 376
  22.3.1 The AGRIT survey 377
  22.3.2 Durum wheat yield forecast 378
22.4 Conclusions 384
References 385

Author Index 389

Subject Index 395