

3 Výběr a metodika dopočtu

3.1 Základní a výběrový soubor

Základní soubor pro Integrované šetření v zemědělství 2023 vychází z Registru farem (46 525 aktivních subjektů) a byl vytvořen na základě prahových hodnot (viz definice zemědělského subjektu v kapitole Úvod).

Z celkem 46 525 aktivních zemědělských subjektů vedených v Registru farem bylo do základního souboru (ZS) vygenerováno 33 745 jednotek. Ty v součtu obhospodařují 99,4 % celkové zemědělské půdy a chovají 98,6 % dobytčích jednotek (DJ).

Výběrový soubor (VS) byl stanoven jako stratifikovaný náhodný výběr ze základního souboru na základě příslušnosti k definované dopočtové skupině (stratu). Každé dopočtové skupině bylo přiřazeno určité procento výběru z celkového počtu jednotek v dané skupině (stratu) v rámci základního souboru. Jednotky s nižším procentem výběru než 100 % byly ze základního souboru generovány náhodným výběrem bez opakování. K tomuto účelu byl využit generátor náhodných čísel v programu SAS Enterprise Guide 4.2.

3.2 Stratifikace a dopočet

Zpravodajské jednotky byly zařazeny do strat, která byla kombinací tří proměnných:

- právní forma
- velikostní kategorie,
- příslušnost ke kraji sídlem farmy nebo bydlištěm fyzické osoby,
- příslušnost k zemědělské výrobní podoblasti.

Kód strata je tvořen sedmiznakovým řetězcem tvaru XAABBCC, kde:

- X nabývá hodnoty 1 pro subjekty fyzických osob a hodnoty 2 pro subjekty právnických osob.
- AA nabývá hodnoty 11 až 15 a 99 a rozlišuje velikost subjektu na základě výměry zemědělské půdy a/nebo dobytčích jednotek:

AA	velikostní kategorie	% výběru ze ZS - FO	% výběru ze ZS - PO
11	9,99 a méně ha zem. půdy nebo 1,99 a méně DJ	8	15
12	10,00 – 24,99 ha zem. půdy nebo 2 – 4,99 DJ	10	25
13	25,00 – 49,99 ha zem. půdy. nebo 5 – 9,99 DJ	15	50
14	50,00 – 99,99 ha zem. půdy nebo 10 – 24,99 DJ	22	50
99	100,00 a více ha zem. půdy	100	100
99	25 a více DJ	100	100
99	40 a více kusů drůbeže	100	100
99	20 a více kusů ovcí	100	100
99	5 a více kusů koz	100	100
99	1,50 a více ha vinic	100	100
99	2,00 a více ha sadů	100	100
99	1,10 a více ha zeleniny	100	100
99	0,30 a více ha květin	100	100
99	0,30 a více ha jahodníku	100	100

Pro zařazení do jednotlivých velikostních kategorií mělo dosažení minimálního počtu kusů v kategorii 99 přednost před vyjádřením prostřednictvím zemědělské půdy či dobytčích jednotek.

- BB je zkrácený kód kraje a nabývá hodnot: 11, 21, 31, 32, 41, 42, 51, 52, 53, 61, 62, 71, 72 a 81.

- CC značí příslušnost k výrobní zemědělské podoblasti:
 - oblast kukuřičná s podoblastmi 11, 12, 13,
 - oblast řepařská s podoblastmi 21, 22, 23,
 - oblast bramborářská s podoblastmi 31, 32,
 - oblast bramborářsko-ovesná s podoblastí 33,
 - oblast horská s podoblastmi 41, 42.

Vzhledem k celkové malé výměře podoblastí 11, 12 a 13 byla jednotkám hospodařícím v podoblasti 12 a 13 přidělena pro výběr i dopočty podoblast 11. Obdobně byla pak nahrazena podoblast 42 podoblastí 41. Přípustné jsou tedy kódy podoblastí 11, 21, 22, 23, 31, 32, 33 a 41.

Počet takto zkonstruovaných dopočtových strat dosáhl hodnoty 928 kombinací, přestože teoreticky bylo možné dosáhnout celkem 1120 (2 x 5 x 14 x 8) kombinací. V tomto členění byl stratifikační kód využit pouze pro generování VS (zachování proporcionality mezi kraji) a pro finální agregaci podle krajů. Při finálním dopočtu na ZS byl využit kód ve tvaru AACCC, tedy maximálně 40 dopočtových strat.

OŠETŘENÍ NON-RESPONSE

Neúplně (částečně) vyplněné výkazy musely být před dopočtovou procedurou vyhledány systémem kontrol a manuálně doplněny individuálním posouzením:

- a) podobným zpravodajským jednotkám v rámci strata se přiřadí zprůměrované hodnoty za ZJ v daném stratu
- b) u zpravodajské jednotky vykazující extrém se imputují poslední známá data.

Jednotky, které odmítly vyplnit výkaz, byly zahrnuty do dopočtových strat podle posledních známých údajů. Ve stratech obsahujících takovéto jednotky byly dopočtové koeficienty navýšeny podle procenta jejich zastoupení.

PŘEPOČET NA ZÁKLADNÍ SOUBOR

Přepočtový koeficient každého strata K_p je dán vztahem:

$$K_p = N_z / n_v, \text{ kde}$$

K_p je přepočtový koeficient

N_z je počet jednotek základního souboru vytvořeného z registru farem za dané stratum

n_v je počet jednotek výběrového souboru za dané stratum

Výstupem dopočtu je agregovaný soubor vzniklý z datového souboru, jehož většinu tvoří agregáty za jednotlivé dopočtové skupiny.

Poznámka: Při výpočtu směrodatné chyby bylo nalezeno 1 718 jednotek s extrémními hodnotami, jejichž extrapolační koeficient byl změněn na hodnotu 1. Tyto jednotky byly přeřazeny do 100% strata.

3.3 Výstupní tabulky

Zpracované tabulky v publikaci sumarizují dopočtené údaje na základní soubor (zemědělství ČR) za zemědělské subjekty, které obhospodařují alespoň 98 % zemědělské půdy nebo chovají alespoň 98 % z celkového počtu dobytčích jednotek.

Publikované výsledky jsou dopočteny z nezaokrouhlených hodnot (data byla dopočítávána s přesností na 6 desetinných míst). Vyčerpávající součet v tabulkách se nemusí vždy přesně rovnat součtu svých složek, tak jak jsou v zaokrouhlené podobě uvedené v tabulkách.

