**Minőségjelentés**

**Nettó fakitermelés (OSAP 1257)**

1. **Általános információk:**

Adatszolgáltatók meghatározása: állami erdőgazdaságok[[1]](#footnote-1) és egyéb erdőgazdálkodók

Kijelölt adatszolgáltatók száma: 545

Tényleges adatszolgáltatók száma: 382

Elektronikus beérkezés száma/aránya: 381 / 100 %

Nullás adatszolgáltatók száma: 260

A kérdőív adathelyeinek száma: 240

Adatgyűjtés módja: kombinált (az állami erdőgazdaságokra teljes körű, az egyéb erdőgazdálkodók esetében mintavételes)

Gyakorisága: éves

Tárgyidőszak: 2020.

Beérkezési határidő: 2021. március 17.

Tervezett publikálás időpontja: 2021.06.30.

1. **Az adatgyűjtés tárgyidőszaki minősége**
	1. **Megbízhatóság**

A korábbi évekhez képest a következő módon módosítottuk a mintavételt: az állami szektort (állami erdőgazdaságok, nemzeti parkok, vízügyi szervek) teljes körűen mintavételeztük, a magánszektorban (beleértve a közösségi szektort is) pedig három réteget különítettünk el az erdőterv alapján várhatóan (egyeztetett sürgősség = 1) fahasználattal érintett terület nagysága szerint (5 hektár ill. azalatti – „kis”; 5 hektárnál nagyobb, de 50 hektárnál kisebb . „közepes”; 50 hektárnál nagyobb – „nagy”). Erre azért volt szükség, mert a korábbiakkal ellentétben a mintavételt súlyozatlan módon végeztük annak érdekében, hogy a magánszektorra vonatkozóan is közvetlenül a mintából lehessen becsülni az országos szintű nettó fakitermelést.

A módosítás eredményeképpen az előző évekhez képest jelentősen megnőtt a nullás adatszolgáltatások száma. Ez elsősorban a kis fahasználati területű erdőgazdálkodókra vezethető vissza. Emellett csökkent a beérkezett kérdőívek száma is, ami feltehetően részben a pandémiás helyzet következménye. A nullás adatszolgáltatások arányának növekedése a csökkent válaszadási hajlandósággal együttesen a statisztikák bizonytalanságát jelentősen növeli. Ez szükségessé teszi a későbbiekben a mintanagyság növelését, amit már jövőre tervezünk végrehajtani.

A beérkezett adatokat azok jóváhagyása előtt tisztítjuk. Ennek során változótípusonként megvizsgáljuk az adatok eloszlását, és statisztikai módszerekkel azonosítjuk az előforduló kiugró értékeket. Ezek leggyakoribb oka a nem megfelelő mértékegység használata, amit az adatszolgáltató megkeresése nélkül is javítani lehet. A kérdéses eseteket az adatszolgáltató megkeresésével tisztázzuk. A 1257-es számú adatgyűjtésben kiugró ill. „gyanús” értékek jellemzően nem fordulnak elő, így ilyeneket az idén sem találtunk.

Statisztikai értelemben az adatok megbízhatóságának két, egymástól független összetevője van: a pontosság (*accuracy*), ami az adatokban lévő torzítottságot, vagyis a szisztematikus hibát fejezi ki, valamint a precizitás (*precision*), ami az adatokban rejlő véletlenszerű hiba számszerűsítése. Mindkét típusú hibának számos oka lehet. Jelen esetben az adatok torzítottsága elsősorban akkor lenne feltételezhető, ha az adatszolgáltatónak anyagi érdeke fűződne ahhoz, hogy a ténylegesen kitermelt fatérfogathoz képest tudatosan többet vagy kevesebbet jelentsen be. Ilyen érdek hiányában az adatokat torzítatlannak tételezhetjük fel.



1. ábra A tűzifa arány átlag megbízhatósága a mintanagyság függvényében a magánszektor „kis” rétegében. A konfidencia intervallumok 95 %-os megbízhatósági szintnek felelnek meg.

A véletlenszerű hibákat elsősorban a mintavételi hiba okozza. A véletlenszerű mintavételi hibát konfidencia intervallum számításával lehet számszerűsíteni. (Az állami szektorban a mintavétel teljes körű, így mintavételezési hiba nincs). A konfidencia intervallum az adott megbízhatósági szintnek megfelelő valószínűséggel fed át[[2]](#footnote-2) a kérdéses statisztika tényleges értékével. A mintavételi hibát a mintanagyság növelésével lehet csökkenteni. A statisztikák megbízhatóságának szempontjából ezért kulcsfontosságú az optimális mintanagyság meghatározása. Bonyolítja a helyzetet, hogy a 1257-es nyilvántartási számú adatgyűjtés sokváltozós. A mintanagyságnak az adatok megbízhatóságára gyakorolt hatását szimulációs eljárással becsülhetjük. Ennek során a beérkezett adatokból veszünk véletlenszerűen különböző nagyságú mintákat, és azokból konfidencia intervallumokat számítunk. Ritkább fafajok ill. választékok esetében az adatok kevésbé megbízhatók (vagyis a konfidencia intervallumok szélesebbek).

Az 1. ábra a magánszektor „kis” rétegében szemlélteti a mintanagyság hatását a becsült tűzifa arány átlag megbízhatóságára fafajtól függetlenül. A tűzifa arány tehát ebben az esetben azt mutatja meg, hogy az összes nettó fakitermelés hányad részét értékesítették tűzifaként. A grafikonról leolvasható, hogy a beérkezett adatokból számolt statisztika erősen bizonytalan (+/- 25 %), ami a mintanagyság növelését feltétlenül indokolttá teszi. Fafajcsoport szinten a konfidencia intervallumok még szélesebbek. (2. ábra).



2. ábra A mintából becsült tűzifatermelés a mintanagyság függvényében a magánszektor „közepes” rétegében. A konfidencia intervallumok 95 %-os megbízhatósági szintnek felelnek meg.

**Indikátorok:**

Egység szintű súlyozatlan beérkezési arány: beérkezett kérdőívek száma/összes adatszolgáltató

70 %

Tétel szintű súlyozatlan beérkezési arány: beérkezett válaszok száma/releváns adatszolgáltatók száma adathelyenként:

70 %

Egység szintű imputálási arány: imputált adatszolgáltatók száma/összes adatszolgáltató

Nem volt szükség imputálásra.

Tétel szintű imputálási arány: imputált értékek száma/összes érték adathelyenként

Nem volt szükség imputálásra.

Hibajavítások száma és aránya:

Javított rekordok száma: 0

Adatszolgáltatók megkeresésével javított hibák aránya: -

Adatfeldolgozó által javított hibák aránya: -

**Időszerűség:** a tervezett közzététel és a leírt esemény referenciaideje közötti idő hossza

6 hónap.

* 1. **Időbeli pontosság:** a publikálás tervezett időpontja tartható-e, ha nem mi ennek az oka (pl. minőség-ellenőrzés, javítás)

Tartható.

* 1. **Hozzáférhetőség:**

**Indikátorok:**

Védett adatok száma és aránya:

Állami erdőgazdaságok: 27 db, 11 %

Nemzeti Parkok: 57 db, 24 %

Vízügyi szervek: 44 db, 18 %

Magánszektor „kis” réteg: 66 db, 28 %

Magánszektor „közepes” réteg: 44 db, 18 %

Magánszektor „nagy” réteg: 44 db, 18 %

* 1. **Összehasonlíthatóság és koherencia:**

**Az előállított statisztikai adatok és velük kapcsolatban álló más adatok összehasonlítása**:

A 1257-es nyilvántartási számú adatlap országos szinten egyedülálló, vagyis nincs olyan másik forrásból elérhető adat, amivel a statisztikák összehasonlíthatók lennének. Az adatlapon belüli elemzések, az állami erdőgazdaságok és az egyéb erdőgazdálkodók adatainak összehasonlítása ugyanakkor informatív, hiszen ezek egymástól függetlennek tekinthetők.

Fontos indikátor a tűzifa fajonkénti aránya (3. ábra), ami az adott fafaj alapvető felhasználásáról (ipari vagy tüzelési cél) ad információt. A fafajok többségénél nincs kiugró különbség a szektorok között. Figyelembe kell venni ugyanakkor az egyes szektorok közötti fafajösszetételbeli különbségeket is (vízügyi szerveknél magas a nyár és a fűz aránya, a magánszektorban az akác a legfontosabb fafaj).

Az iparifa választékok közül a fűrészipari rönk a legnagyobb mennyiségben előállított, az összes kitermelt fatérfogathoz viszonyított aránya szintén fontos jellemzője az adott fafaj hasznosításának. Bár egyes esetekben a szektorok közötti eltérés statisztikailag szignifikánsnak bizonyul, a beérkezett adatok minőségét megkérdőjelező jelenség (pl. egyes szektorok több fafajnál is tapasztalható nem várt irányú kiugróan magas eltérése) nem tapasztalható (4. ábra).

A szektor szerint megbontott adatok összességében jól összehasonlíthatók, koherensek, ami az adatok megbízhatóságát támasztja alá.

**Időbeli összehasonlítás**: a gazdasági, szakpolitikai és egyéb körülmények alapján az adatok időbeli változása elfogadható, magyarázható-e

Az adott évi nettó fakitermelést, választékszerkezetet egyrészt a fahasználati (fakitermelési) lehetőségek, másrészt a piac határozza meg. Egyik évről a másikra hirtelen változást legfeljebb egy váratlan esemény (pl. természeti katasztrófa) okozhat. Ilyen változást az idei évben nem tapasztaltunk. A tavalyi adatokkal való összehasonlítást a statisztikai közleményben adjuk közre.



3. ábra A tűzifa aránya fafajonként és szektoronként. A konfidencia intervallumok 95 %-os megbízhatósági szintnek felelnek meg. Az üresen hagyott oszlopok esetében a mintanagyság nem érte el a hármat.



4. ábra A fűrészrönk aránya fafajonként és szektoronként. A konfidencia intervallumok 95 %-os megbízhatósági szintnek felelnek meg. A hiányzó oszlopok esetében a mintanagyság nem érte el a hármat.

1. A jelentésben közölt statisztikákban az állami erdőgazdaságok magukban foglalják a HM vagyonkezelésében lévő erdőket is. [↑](#footnote-ref-1)
2. A becsülni kívánt statisztika fix értékű, a számított konfidencia intervallum az, ami ahhoz képest eltérhet. Helytelen ezért azt mondani, hogy az intervallum adott valószínűséggel „tartalmazza” a statisztika tényleges értékét. [↑](#footnote-ref-2)