

Minőségjelentés

Erdei faválasztékok értékesítési árai (OSAP 1259/2)

1. Általános információk:

Adatszolgáltatók meghatározása:	állami erdőgazdaságok ¹ és egyéb erdőgazdálkodók
Kijelölt adatszolgáltatók száma:	522
Tényleges adatszolgáltatók száma:	350
Elektronikus beérkezés száma/aránya:	350 / 100 %
Nullás adatszolgáltatók száma:	160
A kérdőív adathelyeinek száma:	31
Adatgyűjtés módja:	kombinált (az állami erdőgazdaságokra teljes körű, az egyéb erdőgazdálkodók esetében mintavételes)
Gyakorisága:	éves
Tárgyidőszak:	2019
Beérkezési határidő:	2020. március 17.
Tervezett publikálás időpontja:	2020.06.30.

2. Az adatgyűjtés tárgyidőszaki minősége

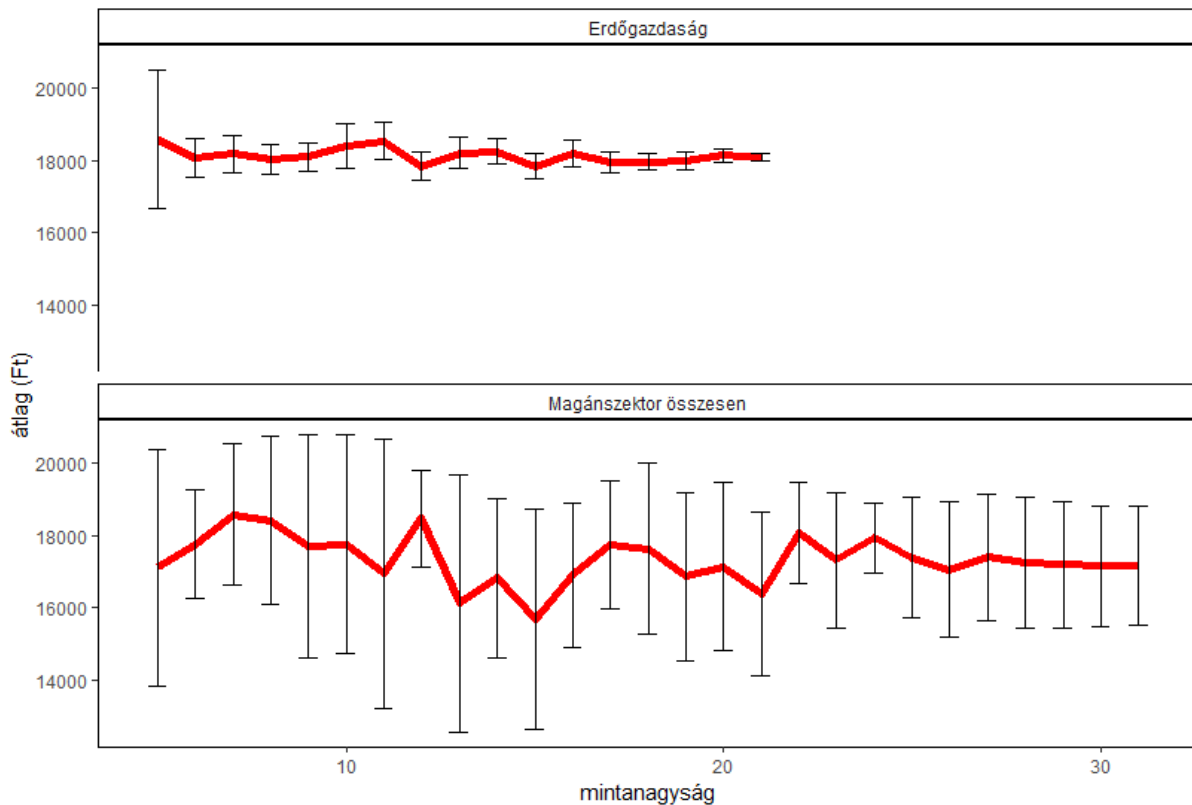
2.1. Megbízhatóság

A beérkezett adatokat azok jóváhagyása előtt tisztítjuk. Ennek során változó típusonként 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.

Statisztikai értelemben az adatok megbízhatóságának két, egymástól független összetevője van: a pontosság (*accuracy*), ami az adatokban lévő torzítottságot, vagyis a szisztematikus

¹ A jelentésben közölt statisztikákban – a védett adatokkal kapcsolatosak kivételével – az állami erdőgazdaságok magukban foglalják a HM vagyongazdálkodó által művelt erdőket is.

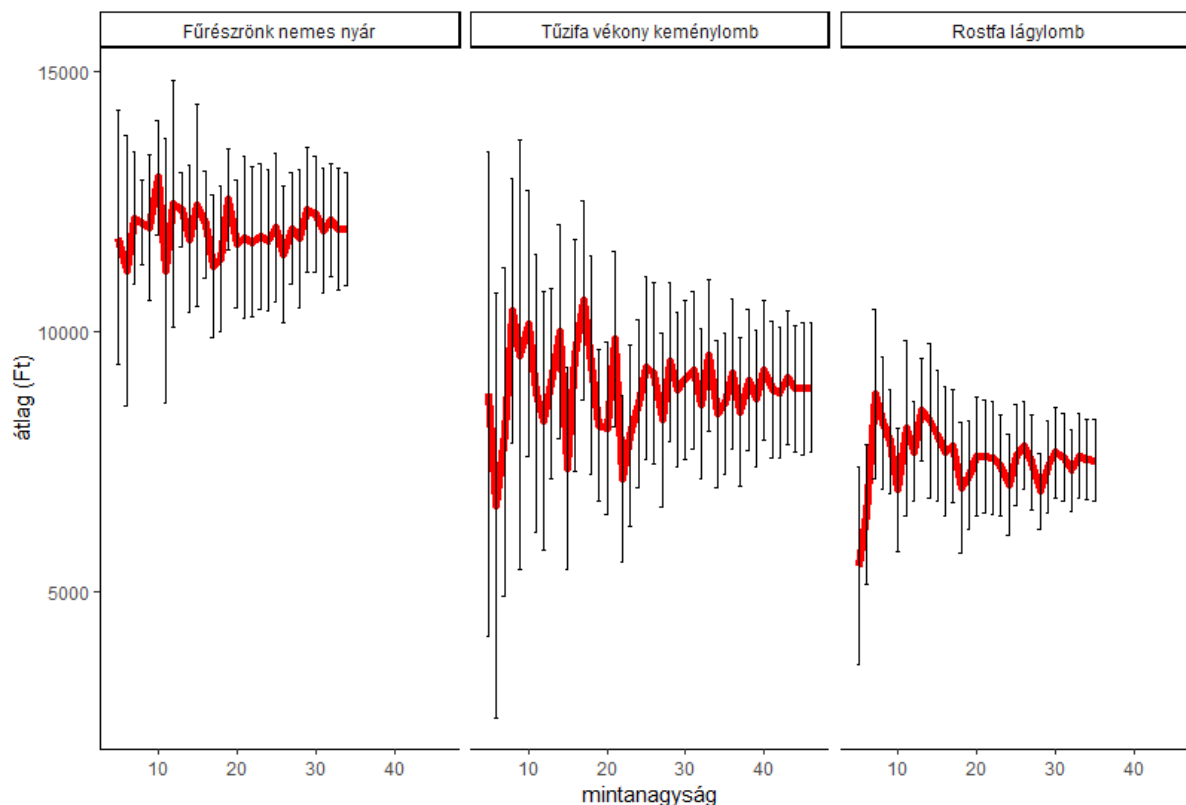
hibát fejezi ki, valamint a precizitás (*precision*), ami az adatokban rejlő véletlenszerű hiba számszerűsítése. Mindkét típusú hibának számos oka lehet. Jelen esetben az adatok torzítottsága elsősorban akkor lenne feltételezhető, ha az adatszolgáltatónak anyagi érdeke fűződne ahhoz, hogy a tényleges árakhoz képest tudatosan többet vagy kevesebbet jelentsen be. Ilyen érdek azonban nincs, hiszen az adatok csak statisztikai célra használhatók fel, hatósági eljárás azok alapján nem indítható.



1. ábra 1 m-es tölgy/cser tűzifa átlagos árának a megbízhatósága a mintanagyság függvényében az erdőgazdaságok ill. a magánszektor esetében. A konfidencia intervallumok 95 %-os megbízhatósági szintnek felelnek meg.

A véletlenszerű hibák főként a mintavételi hibára vezethetők vissza. A véletlenszerű mintavételi hibát konfidencia intervallum számításával lehet számszerűsíteni. (Állami erdőgazdaságoknál a mintavétel teljes körű, így mintavételezési hiba csak abban az esetben léphet fel, ha adott erdőgazdaság adott évre valamilyen oknál fogva nem közöl adott értéket). A konfidencia intervallum a megbízhatósági szintnek megfelelő valószínűséggel fed át² 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 1259-es nyilvántartási számú adatgyűjtés sokváltozós, egyidejűleg több termékről gyűjtünk adatokat.

² A becsléni 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.



2. ábra Egyes fatermékek átlagos árának a megbízhatósága a mintanagyság függvényében a magánszektor esetében. A konfidencia intervallumok 95 %-os megbízhatósági szintnek felelnek meg.

A mintanagyságnak az adatok megbízhatóságára gyakorolt hatását szimulációs eljárással becsülhetjük. Ennek során a beérkezett adatokból veszünk véletlenszerűen különböző nagyságú mintákat, és ezekből konfidencia intervallumokat számítunk. Ritkább fafajok és fatermékek esetében az adatok kevésbé megbízhatók (vagyis a konfidencia intervallumok szélesebbek). Az 1. és a 2. ábraábrán példaként bemutatott fatermékek esetében 30 körüli mintanagyság már elegendőnek látszik megbízható statisztikák nyeréséhez. Efölött a mintanagyság fölött a konfidencia intervallum további szűkülése már nem számottevő, az átlag alig változik. A mintanagyság és a megbízhatóság közötti kapcsolat termékféleségenként és szektoronként némileg eltérő. Összességében megállapítható, hogy az erdőgazdaságok és a magánszektor esetében a mintanagyság elegendő megbízható statisztikák előállításához. Az egyéb állami és a közösségi szektorból azonban a legtöbb termékre nem érkezett kellő mennyiségű adat. Ennek oka, hogy ezekhez a szektorokhoz viszonylag kevés erdőgazdálkodó tartozik, még kevesebb az, aki a tárgyévben fakitermelést végzett. Így az általuk adott évben értékesített termékek köre is igen csekély.

Indikátorok

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

67 %

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

67 %

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/összes rekord: 12/1368

Adatszolgáltatók megkeresésével javított hibák aránya: -

Adatfeldolgozó által javított hibák aránya: 0,9 %

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

6 hónap.

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

Tartható.

2.3. Hozzáférhetőség

Indikátorok:

Védett adatok száma és aránya:

Állami erdőgazdaságok: 0 db, 0 %

HM vagyonkezelő: 15 db, 48 %

Állami szektor: 11 db, 36 %

Közösségi szektor: 12 db, 39 %

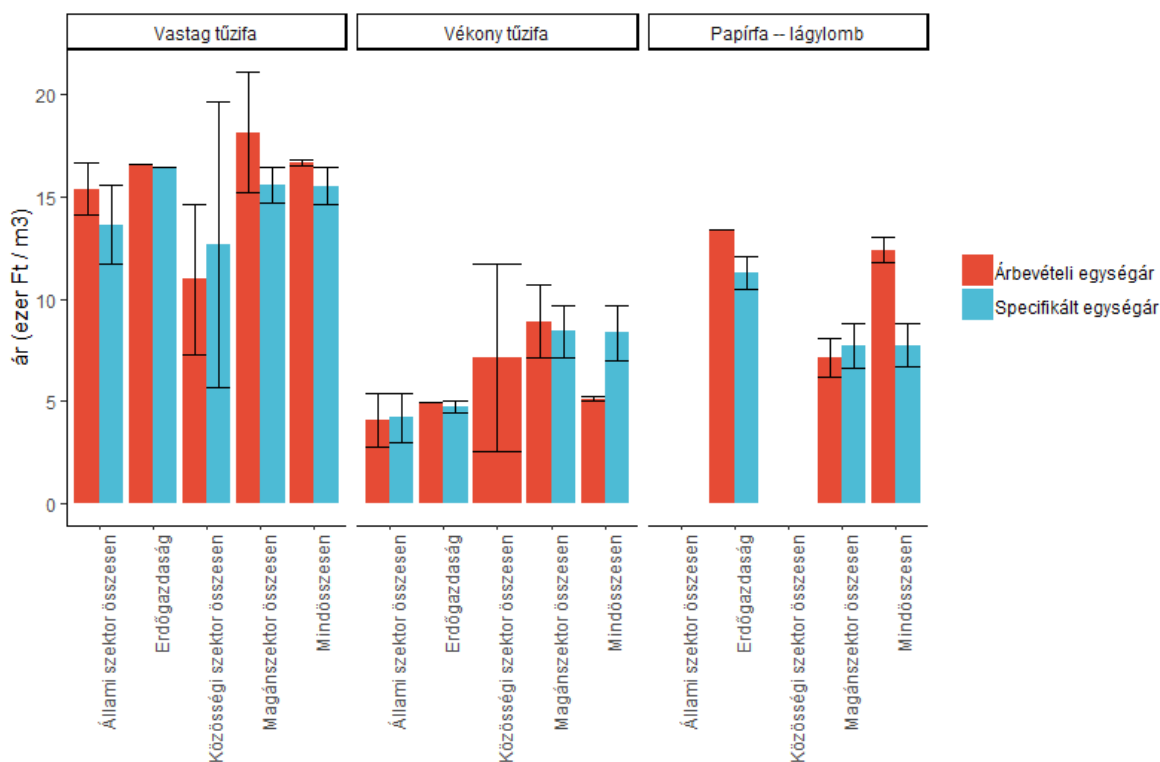
Magánszektor: 1 db, 3 %

2.4. Összehasonlíthatóság és koherencia

Az árstatisztikák a 1259-es számú adatgyűjtés *Erdei fatermékek termelése és készletváltozása* c. résztáblázat árbevételi adataiból számítható egységár statisztikákkal vethetők egybe. Az előbbi táblázat az utóbbtól abban különbözik, hogy az adott fatermék minősége, feldolgozottsága és az értékesítés helyszíne jobban körülhatárolt, valamint nem feltétlenül az adott évi könyvelésen alapul (specifikált egységár). Ennek ellenére a két adatlap árstatisztikái között kiugróan magas eltérésnek nem szabadna lennie. Egyes fatermékekre az összehasonlítást a 3. ábraábrán mutatjuk be. További összehasonlításokat az *Erdei fatermékek termelése és készletváltozása* c. adatgyűjtéshez tartozó minőségjelentésben közlünk. Összességében megállapítható, hogy a két adatlap statisztikai közötti különbségek nem kérdőjelezik meg az adatok megbízhatóságát.

Időbeli összehasonlítás

Az árstatisztikákat az idei év során gyűjtöttük először ebben a formában, így időbeli összehasonlítás nem tehető.



3. ábra A1259-es számú adatgyűjtés két résztáblázatából számolt árak összehasonlítása. A konfidencia intervallumok 95 %-os megbízhatósági szintnek felelnek meg. A háromnál kisebb mintanagysághoz tartozó értékeket nem jelenítettük meg.