



Analyse van de haalbaarheid en de operationaliteit van
een inventarisatie van de leegstaande woningen in het
Brussels Hoofdstedelijk Gewest
Synthese

Rembert De Blander (VUB), Anneloes Vandenbroucke (BSI),
Pierre Marissal & Benjamin Wayens (ULB)

29 februari 2024



Het opzet van deze studie¹ was om na te gaan of het mogelijk is om, door het koppelen van (bestaande) administratieve gegevens, een voorspelling te maken van de kans dat een concrete woning in het Brussels Hoofdstedelijk Gewest (BHG) leegstaat, met het oog op het gericht opsporen van leegstaande woningen. Deze oefening moet begrepen worden in de context van de toenemende nood aan (betaalbare) woningen.

Doorheen de looptijd van het project werden drie rapporten opgeleverd, nl. één per werkjaar. Omdat deze rapporten grotendeels technisch van aard zijn, hebben we daarnaast deze synthese gemaakt. Om de lezer duidelijk te maken wat er zoal kan gevonden worden in de technische rapporten, wordt er regelmatig verwezen naar 'TR 2021' (technisch rapport 1e werkjaar), 'TR 2022' (technisch rapport 2e werkjaar) en 'TR 2024' (technisch rapport 3e werkjaar). Deze rapporten kunnen opgevraagd worden bij Brussel Huisvesting.

1 Beschrijving gebruikte datasets

We bespreken hieronder de verschillende databestanden die gebruikt werden om de leegstandsmodellen te schatten.

Datasets aan de inputzijde (predictoren)

- **Rijksregister (RRN)**: voor de periode 2015 – 2021 alle adressen uit het Brussels Hoofdstedelijk Gewest (BHG) waar in de loop van het betrokken jaar minstens 1 persoon gedomicilieerd was. We stellen een gestage toename vast van het aantal adressen waar minstens één persoon gedomicilieerd is (van zo'n 434 000 in 2015 tot 476 000 in 2021).
- **Kadaster**: voor de periode 2015 – 2021 een typering (aantal en aard) van het vastgoed per perceel in het BHG op 1 januari van het betrokken jaar. Ook voor het kadaster stellen we over de periode 2016 – 2021 een gestage toename vast van het aantal bewoonbare eenheden (van zo'n 507 000 in 2016 tot 537 000 in 2021).
- **Vivaqua**: voor de periode 2015 – 2021 de adressen van abonnees op ±31 december van het betrokken jaar, het aantal woningen en meters op dat adres en het gemiddeld verbruik indien dit „laag” is, anders blijft deze variabele leeg. Het aantal adressen kende ook hier een stijging over de betrokken periode (van zo'n 335.000 in 2015 tot 378 000 in 2021). Het percentage adressen met laag verbruik fluctueert tijdens de betrokken periode tussen de 6.4% en de 8.1% en is vooral sinds 2019 toegenomen.
- **Kruispuntbank van Belgische ondernemingen (KBO)**: voor de jaren 2015 – 2021 de bedrijfsvestigingen in het BHG (ook weer toenemend van zo'n 136.000

¹in opdracht van Brussel Huisvesting

in 2015 tot 203 000 in 2021). Aan de hand van oprichtingsdatum en datum van stopzetting werd een panel van bedrijfsvestigingen in het BHG geconstrueerd.

- **Tweede verblijfstaks:** voor het jaar 2020 de adressen waarvoor een tweede verblijfstaks betaald werd. Dit aantal varieert sterk per gemeente. Zo waren er in het jaar 2020 tien Brusselse gemeenten waar er voor geen enkel adres een belasting tweede verblijf betaald werd (o.a. Brussel-Stad, Ukkel, Anderlecht...). Anderzijds werd er in Sint-Pieters-Woluwe bijvoorbeeld een belasting tweede verblijf gegeven op 226 adressen.

De gegevensbestanden met bouw-aanvragen („NOVA”) en met geregistreerde huurcontracten („MyRent”) werden uiteindelijk niet gebruikt omdat uit eerder onderzoek bleek dat de betrouwbaarheid van beide niet zo groot is (zie TR 2021 voor verdere toelichting).

Datasets aan de outputzijde (trainingssets)

Het eerste jaar van het project werd noodgedwongen gewerkt met een sub-optimale trainingsset. De resultaten van deze eerste modellen worden hier verder niet besproken.

Op verzoek van het onderzoeksteam werd na het eerste jaar, door de administratie, een random steekproef van 5 000 adressen gecontroleerd op het terrein met betrekking tot hun leegstandstoestand. De resulterende dataset vormde de finale trainingsset voor de leegstandsmodellen. Het baseren van de modellen op deze steekproef maakte accuratere schattingen van de kans op leegstand mogelijk.

Steekproef

De **willekeurige steekproef** werd op **gestratificeerde** wijze getrokken. Deze werkwijze geeft zekerheid dat alle types van gebouwen voldoende aanwezig zijn in de steekproef. Er werd concreet gestratificeerd naar gemeente, type huizenblok en type gebouw. De huizenblok-typologie werd gebaseerd op 5 groepen variabelen: woningdensiteit en functies, bereikbaarheid met openbaar vervoer, type gebouw, situatie van de gebruiker (eigenaar, huursituatie...) en kenmerken van de huishoudens. Deze classificatie resulteert in 14 types huizenblokken. Daarbovenop werden de individuele gebouwen onderverdeeld in 5 types (gesloten, (half-)open, flatgebouw, handel, ander). De kruising van beide typologieën met de 19 Brusselse gemeentes levert, na verwijdering van de lege categorieën, een categorisering van de adressen in 1 070 mogelijke strata op. Voor verdere details over de constructie van de steekproef verwijzen we naar TR 2022.

Terreinwerk

De administratie voerde voor alle gebouwen in de random steekproef volgende controles uit: visuele inspectie van de buitenkant van het gebouw, nagaan of er namen op de bel(len) of parlofoons en brievenbussen staan, aanbellen ter controle van aanwezigheid, indien mogelijk de bewoning navragen bij de burens, een uitgebreidere controle van de inkomhal en de gevel. In een volgende stap werd er op kantoor bijkomend een vergelijking gemaakt van het aantal gezinshoofden (rijksregister) met het aantal woon-eenheden (kadaster).

2 Werkwijze koppelen datasets

Omdat België, in tegenstelling tot bijvoorbeeld de Scandinavische landen, geen systeem kent waarbij aan elke woning een unieke woningidentificator wordt toegekend, gebruikt elk databestand dat informatie bevat over wooneenheden een eigen identificator, in de praktijk (onder meer) het adres. De manier waarop deze adressen echter in de bestanden voorkomen verschilt sterk, met als gevolg dat het koppelen van de verschillende bestanden een enorm werk is. Als algemene strategie werd gekozen om de beschikbare datasets in eerste instantie maximaal te koppelen aan BeSt Address². Dit lukte grotendeels tot op het niveau van het gebouw (unieke combinatie van gemeente, straat, huisnummer en huisnummersuffix (A, B, ...)), maar het niveau van de woning bleek door de veelvormigheid van de gebruikte postbus-formats onhaalbaar. Daarmee werden de gebouwen dus onze analyse-eenheid. Ondanks alle ingezette middelen konden in elke dataset een beperkt aantal adressen niet gekoppeld worden (gaande van 0.3% voor KBO en Rijksregister, tot 4.25% voor Vivaqua). Voor een gedetailleerde beschrijving van het „opkuiswerk” verwijzen we naar TR 2021. Alle records die succesvol konden gekoppeld worden aan BeST Add werden vervolgens gekoppeld per wooneenheid en alle gekoppelde records werden finaal gecombineerd op het gebouw-niveau.

Alle software die nodig is om de verschillende datasets aan elkaar te koppelen, inclusief handleiding, werd door het onderzoeksteam bezorgd aan Brussel Huisvesting. De software werd zodanig geconcipeerd dat toekomstige updates van de resulterende gekoppelde database op (semi-)automatische wijze zouden kunnen verlopen.

3 Modellen en methodes

Gezien de terreincontroles werden uitgevoerd in de loop van 2022 werden de modellen geschat op de gekoppelde gegevensbestanden voor de periode 2015 – 21 (we beperken ons hier dus tot verklarende variabelen geobserveerd vóór de uitkomstvariabelen). Op

²<https://bosa.belgium.be/nl/services/best-address-services>

basis van eerdere exploraties werd er voor gekozen om aparte modellen te schatten voor enkelvoudige (ééngezinswoningen) en meervoudige gebouwen (meergezinswoningen).

Afhankelijke variabelen - verschillende definities van leegstand

Er werden modellen gedraaid voor vier afhankelijke variabelen:

1. Leegstand op basis van de ruwe terreingegevens, enkel gebaseerd op het aantal bellen en brievenbussen met en zonder naamsvermelding.

Omdat deze modellen tot een duidelijke overschatting van de leegstand leiden (zie TR 2024 voor meer details), worden ze hier niet verder besproken.

2. Door de administratie gerapporteerde leegstand (R), gebaseerd op de veldwerknooties, aangevuld met een controle van het rijksregister.

Op basis van de verschillende controles werden alle gebouwen geëlassificeerd als „leegstand” (wooneenheden die (quasi-)zeker leegstaan), „partiële leegstand” (sommige maar niet alle wooneenheden in het gebouw staan leeg), „potentiële leegstand” (geen/minder gezinshoofden dan verwacht en visuele controle onduidelijk) en „bevoond”.

Op basis hiervan definieerden we de volgende twee leegstandsvariabelen:

- breed gedefinieerde leegstand (RB) omvat de gebouwen in de categorieën „leegstand”, „partiële leegstand” en „potentiële leegstand”. 314 observaties (6.35% van de steekproef) voldoen hier aan.
- nauw gedefinieerde leegstand (RS) omvat de gebouwen in de categorieën „leegstand” en „partiële leegstand”. (101 gebouwen – 2.04%).

3. Leegstand op basis van doorgedreven administratieve Controles (C), die bestonden uit twee etappes. Na de eerste etappe kregen woningen het statuut „Sans suite” [geen opvolging], „Soupçon d’infraction” [vermoeden van inbreuk] of „Standby”; na de tweede etappe werd het resultaat genoteerd als „Amende” [boete] of „Fin de procédure” [procedure stopgezet] (met ook observaties [Missing]). Aangezien er meerdere dossiers per gebouw kunnen zijn, werden de gegevens geaggregeerd op niveau van het gebouw. Op basis hiervan (*minstens één dossier leverde een boete op / werd gevlagd als „vermoeden van” / werd genoteerd als „Standby”*) definieerden we opnieuw twee leegstandsvariabelen:

- breed gedefinieerde leegstand (CB) omvat alle gebouwen die minstens 1 woning omvatten die een boete gekregen heeft, waarvoor een vermoeden van leegstand bestaat of die „Standby” is (in beide laatste gevallen, gecombineerd met ontbrekende info voor stap 2)

- nauw gedefinieerde leegstand (CS) omvat enkel de gebouwen die een woning omvatten die een boete gekregen heeft
4. Leegstand op basis van bijkomende terreinchecks door IGEAT.
- Een klein deel van de steekproef (zo'n 5%) waarvoor de gerapporteerde classificatie (zie punt 2) vragen opwierp, werd door IGEAT gehercontroleerd op het terrein. Doel was om beter zicht te krijgen op de omstandigheden die soms leiden tot foutief gerapporteerde aantallen woon-eenheden in een gebouw, om atypische omstandigheden te verifiëren en om enkele hypothesen na te gaan i.v.m. criteria die de administratie hanteerde bij het classificeren van haar veldwerk (zie punt 2). Deze hercontroles leverden een aantal interessante bevindingen op (zie TR 2024). Omdat ze echter zo partieel waren en het resultaat van deze modellen niet noemenswaardig verschilde van de modellen voor punt 2, worden ze hier niet verder besproken.

Onafhankelijke variabelen - leegstand 'verklaren'

Op basis van de vijf eerder beschreven datasets, werden verschillende types indicatoren (enkelvoudige, longitudinale en combinatie-indicatoren) geconstrueerd die een bijdrage kunnen leveren aan het voorspellen van leegstand:

- 6 lage waterconsumptie-indicatoren van de vorm 'Gebouw bevat minstens één woning met lage waterconsumptie in het jaar 2021', '... gedurende de hele periode 2020 – 2021', '... gedurende de hele periode 2019 – 2021', ...
- 6 domiciliëringsindicatoren van de vorm 'Gebouw bevat minstens één woning zonder domiciliëring in het jaar 2021', '... gedurende de hele periode 2020 – 2021', '...gedurende de hele periode 2019 – 2021', ...
- Het aantal in het gebouw gedomicilieerde gezinnen (volgens het rijksregister)
- 11 bedrijfsindicatoren:
 - 'Gebouw bevat minstens één bedrijf'
 - 5 indicatoren van de vorm 'Gebouw bevat nu minstens één bedrijf dat opgericht werd in het jaar 2021', '... in de periode 2020 – 2021', '...in de periode 2019 – 2021'...
 - 5 indicatoren van de vorm 'Gebouw bevatte eerder minstens één bedrijf gestopt in het jaar 2021'; '... in de periode 2020 – 2021', '...in de periode 2019 – 2021'...
- Belasting op tweede verblijven: 'Gebouw bevat minstens één woning waarop een belasting tweede verblijf werd geheven in 2020'

- Transitionele ('voorbijgaande') leegstand

Omdat transitionele leegstand kan bijdragen aan een „indruk van leegstand” (zie 'Leegstand op basis van de ruwe terreingegevens'), maar mogelijk net een negatieve invloed heeft op de strengste definitie van leegstand, namelijk de gebouwen die een boete gekregen hebben (CS), hebben we in functie van de modellering een maatstaf voor transitionele leegstand, ofwel turbulentie, geconstrueerd (zie TR 2024). Uiteindelijk blijkt deze maat in slechts één model significant van nul te verschillen, namelijk dat voor de nauwe leegstands-indicator op basis van de doorgedreven administratieve controles (CS), voor enkelvoudige gebouwen ($p < .05$).

- 3 combinatie-indicatoren door het vergelijken van het aantal wooneenheden in verschillende datasets

Aangezien de eenheden in de verschillende gegevensbronnen niet altijd compatibel zijn, werd de nadruk gelegd op de longitudinale indicatoren.

Leegstandsmodellen

Voor elk van de leegstandsvariabelen, werd de kans op leegstand geschat aan de hand van een apart discrete keuze-model (probit) model voor enkelvoudige en meervoudige gebouwen (7×2 modellen in totaal). Zoals hoger aangegeven beperken we ons hier tot de acht modellen (gerapporteerd (R) / doorgecontroleerd (C) x brede (B) / smalle (S) definitie x enkelvoudige / meervoudige gebouwen) die ons het relevantst lijken.

In het algemeen kunnen we stellen dat de modellen geschat voor de enkelvoudige gebouwen (ééngeswoningen) beter presteren dan deze voor de meervoudige gebouwen: de pseudo- R^2 ligt in het eerste geval rond de 0.31 (uitzondering: een pseudo- R^2 van 0.57 voor het CS model), terwijl die in het tweede geval ongeveer 0.22 bedraagt (uitzondering: een pseudo- R^2 van 0.14 voor het CS model). Een tweede vaststelling is dat de variabele 'Gebouw bevat minstens één woning zonder domiciliëring gedurende de hele periode 2015 – 2021' de belangrijkste predictor is in zes van de acht beschouwde modellen, zowel wat betreft grootte van de coëfficiënt als p -waarde (enkel voor de twee C modellen voor enkelvoudige gebouwen, speelt deze variabele niet echt). Voor de enkelvoudige gebouwen is de tweede belangrijkste verklarende variabele in de RB en RS modellen dat er zich in 2021 (het meest recente jaar in het model) in het gebouw minstens één wooneenheid bevond met een lage waterconsumptie. Bij de enkelvoudige CB en CS modellen zijn de belangrijkste verklarende variabelen voor leegstand, dat er een bedrijf gevestigd is in de ééngeswoning (een oneigenlijk gebruik van de woning) en dat er de laatste drie jaar niemand gedomicilieerd was.

Voor de meervoudige gebouwen wordt de breed gedefinieerde leegstand (RB en CB modellen) verder voorspeld door een overtal aan woon-eenheden ($Vivaqua > RRN$) in combinatie met de aanwezigheid van een bedrijf en de vaststelling dat het gebouw minstens één wooneenheid bevat zonder domiciliëring gedurende de periode 2020 – 2021. De aanwezigheid van een wooneenheid zonder domiciliëring gedurende de periode 2019 – 2021 heeft vreemd genoeg een negatief effect op de kans op leegstand. In het RS model verklaart - behalve het al vermeldde 'Gebouw bevat minstens één woning zonder domiciliëring gedurende de hele periode 2015 – 2021' - o.a. ook de indicator 'Gebouw bevat nu minstens één bedrijf dat opgericht werd in de periode 2020 – 2021' mee de kans op leegstand.

Aangezien we op voorhand niet konden uitsluiten dat de leegstand ook beïnvloed wordt door gemeente en huizenblok-type (bovenop de al beschouwde variabelen), hebben we voor elk model ook nagegaan of het opnemen van gemeente-dummies of huizenbloktypedummies de modellen significant verbeterde, maar dat bleek meestal niet het geval te zijn. Voor meer details over de schattingsprocedure en de resultaten van alle geschatte modellen, verwijzen we naar TR 2024.

Leegstandsdetectietool

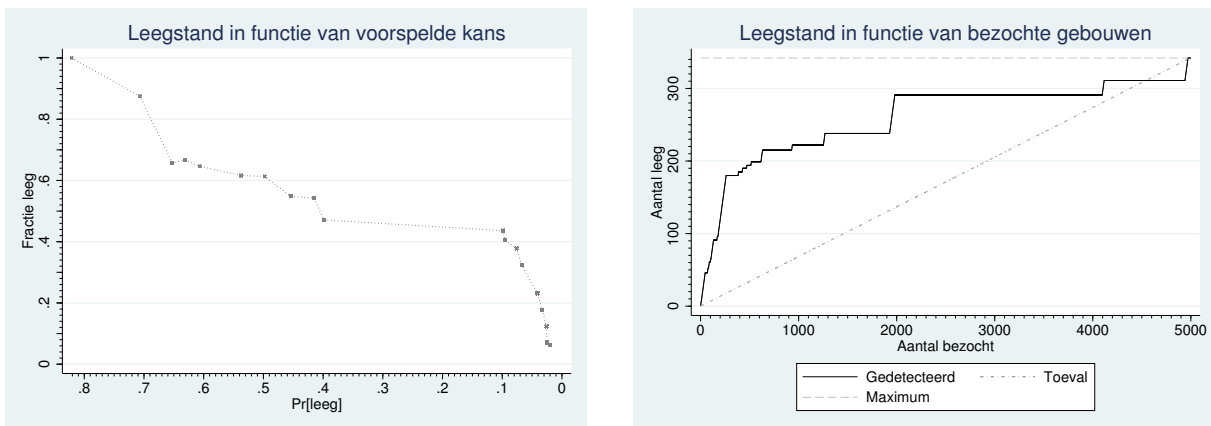
De voorgestelde modellen vormen nu de basis voor de leegstandsdetectietool: door de meest geschikte modellen (één voor enkelvoudige gebouwen en één voor meervoudige gebouwen) toe te passen op alle gebouwen in de gekoppelde dataset (\sim het Brusselse woningenpark), kunnen we een theoretische kans op leegstand berekenen voor elk gebouw. Vervolgens voegen we de kansen voor enkelvoudige en meervoudige gebouwen samen in één variabele en ordenen we alle woningen op basis hiervan, van hoog naar laag. Wanneer woningen nu op leegstand gecontroleerd worden volgens deze rangorde, zou dit een duidelijke efficiëntiewinst moeten opleveren (ten opzichte van willekeurige, 'blinde' controles).

Om de prestaties van de voorgestelde leegstandstool in kaart te brengen, maken we voor de ± 5000 gebouwen in de willekeurige steekproef een vergelijking tussen de berekende, 'theoretische' kans op leegstand die het model voorspelt en de geobserveerde realiteit. Het resultaat presenteren we hieronder in twee soorten grafieken (voorbeelden voor de RB en RS modellen):

- De eerste grafiek geeft telkens de fractie werkelijk leegstaande gebouwen weer, wanneer we alle gebouwen zouden controleren waarvoor de leegstandsindicator groter is dan een bepaalde waarde
- De tweede grafiek geeft het aantal gebouwen met leegstand weer in functie van het aantal gebouwen dat wordt gecontroleerd in volgorde van leegstandsrisico (van hoog naar laag)

Figuur 1 toont het resultaat voor de breed gedefinieerde, gerapporteerde leegstand (RB). De eerste figuur toont aan dat zo'n 100% van de gebouwen leegstaan, wanneer we enkel gebouwen controleren waarvoor het geschatte model een kans van 82% aangeeft (merk op dat de voorspelde kans op leegstand gerangschikt werd van hoog naar laag). Wanneer we nu steeds meer gebouwen (met een lagere voorspelde kans) controleren, dan staan bijvoorbeeld 88% van de gebouwen leeg bij een cut-off waarde van 71% voor de voorspelde kans op leegstand, en eindigen we finaal met de 6.85% leegstand volgens de brede definitie die in de volledige steekproef voorkomt.

In de tweede figuur blijkt de meerwaarde van het model uit de mate waarin het aantal geobserveerde gebouwen met leegstand hoger ligt dan men op basis van toeval mag verwachten. Op basis van toeval detecteren we 6.85% brede leegstand, wat overeenkomt met de punten op de diagonaal tussen de punten (0,0) en (4995,342). In concreto zou een willekeurige controle van 200 gebouwen leiden tot de identificatie van zo'n 13.7 leegstaande gebouwen (in brede zin), terwijl dit aantal met behulp van de tool rond de 180 ligt.

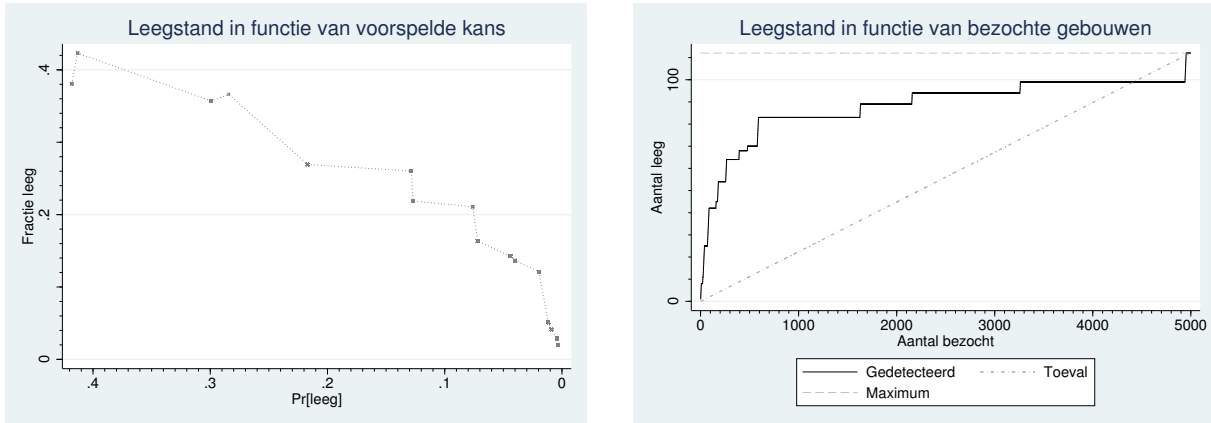


Figuur 1: Prestatie van de voorgestelde detectietool voor de breed gedefinieerde, gerapporteerde leegstand

Figuur 2 geeft beide grafieken weer voor de nauw gedefinieerde, gerapporteerde leegstand (RS). De eerste figuur toont aan dat zo'n 44% van de gebouwen leegstaan, wanneer we enkel gebouwen controleren waarvoor het model een kans van 42% aangeeft (0.58 op de X-as, omdat we van hoog naar laag rangschikken). Wanneer we nu steeds meer gebouwen (met een lagere voorspelde kans) controleren, eindigen we finaal met de 2.24% leegstand volgens de nauwe definitie, het percentage dat in de volledige steekproef voorkomt.

In de tweede figuur blijkt de meerwaarde van de tool opnieuw uit de mate waarin het aantal geobserveerde gebouwen met leegstand hoger ligt dan men op basis van toeval mag verwachten. Op basis van toeval detecteren we 2.24% nauwe leegstand, wat overeenkomt met de punten op de diagonaal tussen de punten (0,0) en (4995,112). In concreto zou

een willekeurige controle van 600 gebouwen leiden tot de identificatie van zo'n 13.44 leegstaande gebouwen (in nauwe zin), terwijl dit aantal met behulp van de tool rond de 85 ligt.



Figuur 2: Prestatie van de voorgestelde detectietool voor de nauw gedefinieerde, gerapporteerde leegstand

Schatting van de leegstand in het BHG

We kunnen nu tenslotte ook, op twee manieren, schattingen maken van het aantal leegstaande woningen in het BHG:

1. Op basis van de willekeurige steekproef, naargelang de gehanteerde definitie van leegstand.
2. Door voor alle gebouwen in de gekoppelde database de voorspelde kans op leegstand te berekenen, opnieuw naargelang de gehanteerde definitie van leegstand

We presenteren hieronder de schattingen op basis van de R en C modellen en maken telkens een aparte berekening voor enkelvoudige en meervoudige gebouwen die we vervolgens combineren tot een totaal. Omwille van de verschillende mogelijke definities (breed/smal) en - in geval van meervoudige gebouwen - de onzekerheid over het aantal leegstaande eenheden per gebouw, geven we voor elk model een onder- en bovengrens aan wat betreft % leegstand en de daarbij horende 'vork' aan leegstaande woningen. Voor meer details over de concrete berekeningswijze verwijzen we naar 'TR 2024'.

Schatting gebaseerd op de willekeurige steekproef

In tabel 1 presenteren we de geschatte onder- en bovengrenzen wat betreft het % leegstand, op basis van de random steekproef. Het combineren van de enkelvoudige en meervoudige gebouwen leidt, als we ons baseren op de gerapporteerde leegstand, tot een geschatte vork van 0.65% à 4.63% leegstaande woningen. Op basis van de doorgecontroleerde (strafbare)

leegstandsmaat krijgen we een geschatte vork van 0.15% à 1.78% leegstaande woningen .

Type	Enkelvoudige		Meervoudige		Totaal	
	Onder	Boven	Onder	Boven	Onder	Boven
Rapportering	1.35	5.96	0.48	4.30	0.65	4.63
Admin. controles	0.35	2.03	0.11	1.72	0.15	1.78

Tabel 1: Vorken van leegstandspercentages voor wooneenheden, per type gebouw en totaal

Om deze geschatte leegstandspercentages om te zetten naar aantal leegstaande wooneenheden, baseren we ons op het aantal wooneenheden in het gebouw volgens de Vivaqua gegevens. In totaal gaat het hier om 77 323 enkelvoudige gebouwen en 86 335 meervoudige gebouwen, waarin zich 429 084 wooneenheden bevinden. Dat brengt het totaal aantal wooneenheden in het BHG op 506 407.

Type	Enkelvoudige		Meervoudige		Totaal	
	Onder	Boven	Onder	Boven	Onder	Boven
Rapportering	1 042	4 606	2 045	18 454	3 087	23 060
Admin. controles	267	1 569	453	7 392	720	8 961

Tabel 2: Vorken van leegstaande wooneenheden in het BHG, per type gebouw en totaal

Tabel 2 presenteert het geschatte aantal leegstaande wooneenheden in het BHG, op basis van de random steekproef. We bekomen dit door het aantal wooneenheden per gebouw te vermenigvuldigen met de type-specifieke leegstandspercentages voor wooneenheden uit Tabel 1. Op basis van de initieel gerapporteerde leegstand, leidt het combineren van de verschillende types gebouwen tot een vork van 3 087 à 23 060 leegstaande woningen. Op basis van de maat voor doorgecontroleerde (strafbare) leegstand krijgen we een geschatte vork van 720 à 8 961 leegstaande woningen .

Schatting gebaseerd op de modellen

We kunnen de schattingen van de leegstandspercentages (Tabel 3) en -aantallen (Tabel 4) nu ook baseren op de geschatte leegstandskansen voor de volledige gebouwendatabase. Uiteraard geeft dit andere concrete aantallen, maar kwalitatief komen de gegevens overeen met deze berekend op de willekeurige steekproef.

Type	Enkelvoudige		Meervoudige		Totaal	
	Onder	Boven	Onder	Boven	Onder	Boven
Rapportering	3.09	10.17	2.06	5.49	2.57	7.80
Admin. controles	0.43	3.25	0.39	2.06	0.41	2.64

Tabel 3: Vorken van leegstandspercentages voor wooneenheden, per type gebouw en totaal

Type	Enkelvoudige		Meervoudige		Totaal	
	Onder	Boven	Onder	Boven	Onder	Boven
Rapportering	1 614	5 807	1 958	23 284	3 572	29 091
Admin. controles	280	2 165	363	9 333	643	11 498

Tabel 4: Vorken van leegstaande wooneenheden in het BHG, per type gebouw en totaal

4 Conclusie en aanbevelingen

Conclusie

Dit onderzoek had tot doel een „Analyse van de haalbaarheid en de operationaliteit van een inventarisatie van de leegstaande woningen in het Brussels Hoofdstedelijk Gewest” uit te voeren. De globale conclusie is dat het inderdaad haalbaar blijkt om door het koppelen van (bestaande) administratieve gegevens, een voorspelling te maken van de kans dat een concrete woning in het Brussels Hoofdstedelijk Gewest leegstaat, met het oog op het gericht opsporen van leegstaande woningen. Concreet leverde het koppelen van 4 databases (m.n. het kadaster, het rijksregister, de Kruispuntbank van Belgische Ondernemingen en gegevens over laag waterverbruik (Vivaqua)) de basis voor een bruikbaar model.

Globaal stellen we vast dat de combinatie-indicator ‘Gebouw bevat minstens één woning zonder domiciliëring gedurende de hele periode 2015–2021’ de sterkste voorspellende waarde heeft. In bijna elk model zit ook een significante longitudinale indicator, wat het belang aangeeft van het kunnen koppelen van de verschillende jaargangen van een dataset (constructie panel) als men leegstand wil kunnen voorspellen.

Een handig bijproduct van het succesvol schatten van een reeks leegstandsmodellen is de ontwikkeling van een leegstandsdetectietool die toelaat het risico op leegstand te voorspellen. Dit laat de betrokken administratie toe veel gericht te werk te gaan bij het opsporen van leegstaande woningen.

Tot slot is deze studie, voor zover we weten, de eerste die het niveau van leegstand in het Brussels Hoofdstedelijk Gewest schat (onder de vorm van leegstandsvorken) op basis van een willekeurige steekproef. Het veldwerk dat door de administratie werd verricht om de leegstandsstatus te verifiëren van de gebouwen in de steekproef, was hierbij cruciaal. Zonder dit veldwerk zouden dergelijke leegstandsschattingen niet mogelijk zijn geweest en zou de leegstandsdetectietool minder nauwkeurig zijn geweest. Om de werking van de tool te blijven garanderen, dient periodiek een nieuwe willekeurige steekproef en bijhorend veldwerk uitgevoerd te worden. Even essentieel voor het welslagen van toekomstige updates is het dat - gegeven de bijdrage van de longitudinale indicatoren - de verschillende data-leveranciers zich aan een vast dataformaat houden.

Aanbevelingen

- Het is aan te raden de *succes-rate* van de controles die Brussel Huisvesting uitvoert te monitoren, zodat we naderhand een vergelijking kunnen maken tussen de voorspelde kans op leegstand en de reële situatie. Dit zal ook toelaten na te gaan welke van de leegstandsschattingen het meest betrouwbaar is.
- De analyses suggereren dat wat soms wordt geïnterpreteerd als een leegstaande woning, in feite een woning is zonder domicilie of een woning die gebruikt wordt voor niet-residentiële doeleinden. Het lijkt daarom evenzeer belangrijk **het 'juiste' gebruik van woningen** te stimuleren, onder andere door toeristische accommodatie beter te reguleren en door alternatieve locaties aan te bieden voor niet-residentieel gebruik.
- Verschillende indicatoren en de laatst beschikbare cijfers (uit de socio-economische enquête van 2001) wijzen op het naast elkaar bestaan van **over- en onderbezetting**. Een betere inventarisatie van het woningenpark zou hier meer klaarheid in kunnen scheppen. Onderbezetting zou vervolgens gericht verminderd kunnen worden door bijvoorbeeld een groter aanbod te creëren dat aangepast is aan alleenstaanden en koppels wiens kinderen het ouderlijk huis verlaten hebben.
- De lage transitionele (tijdelijke, voorbijgaande) leegstand wijst op het belang van het creëren van een bijkomend aanbod van betaalbare woningen.
- Een **unieke woningidentificator** (of minstens een uniforme adrescodering) zou niet alleen de koppeling van verschillende gegevensbanken vergemakkelijken, wat de performantie van de modellen zou kunnen verbeteren, maar bovendien een correcte inventarisatie van het woningenpark mogelijk maken. Dit laatste vergemakkelijkt verschillende objectieven: het monitoren van de leegstand wordt eenvoudiger én het biedt een correcte belastbare basis voor de onroerende voorheffing. Een unieke woningidentificator zou bovendien toelaten de in- en uitschrijvingsprocedures in het rijksregister te stroomlijnen: het creëren van pseudo-woningen op hetzelfde adres kan hierdoor vermeden worden, schrappingen kunnen beter opgevolgd worden, huishoudens worden eenduidiger gedefinieerd. Bovendien zou een eenvoudige vergelijking van de unieke woningidentificatoren met het rijksregister volstaan om de nader te controleren woningen te identificeren.
- Zolang er geen unieke woning-identificator ingevoerd wordt, moet de effectieve toepassing van **BeSt Add** bij alle Brusselse actoren, privé en publiek, een prioriteit zijn als we het kruisen van bestaande gegevens willen vergemakkelijken en versnellen en de kwaliteit ervan willen verbeteren.