

## Kvalitetskontrol af FLAT-adresser

<b>Forfatter:</b>	Jesper Stenstrup & Anette Vainer
<b>Godkendt af:</b>	Nils Bo Wille-Jørgensen
<b>Version:</b>	1.1
<b>Projektnr.:</b>	50.1434.00
<b>Filnavn:</b>	S:\50_1400\501434\jss\Doc\rapport97.doc
<b>Antal sider:</b>	23
<b>Dato:</b>	2000-03-07
<b>Kunde:</b>	Kort & Matrikelstyrelsen Rentemestervej 8 2400 København N

## VERSIONSHISTORIE

Version	Af	Dato	Ændringsbeskrivelse
0.1	JsS	2000-01-28	Første version.
0.2	JsS	2000-02-23	under revision
1.0	JsS	2000-03-07	Endelig version
1.1	AeV	2000-04-07	Grafer for fordeling af afstande udskiftet

## INDHOLDSFORTEGNELSE

<b>1. INDLEDNING</b> .....	<b>4</b>
<b>2. RESUME</b> .....	<b>4</b>
<b>3. KVALITETSBESKRIVELSENS GRUNDLAG</b> .....	<b>5</b>
<b>3.1 Kvalitetsmålsætning for FLAT adresser</b> .....	<b>5</b>
3.1.1 Målsætning 1: Fuldstændighed af adressemængde .....	5
3.1.2 Målsætning 2: Fuldstændighed af stedfæstelse .....	6
3.1.3 Målsætning 3: Beliggenhed på korrekt matrikel lod .....	6
3.1.4 Målsætning 4: Afstand til den korrekt placerede adresse for adresser, der er placeret på korrekt matrikelod .....	6
<b>4. VURDERING AF PRODUKTIONSMETODEN I FORHOLD TIL MÅLSÆTNINGERNE</b> ...	<b>6</b>
<b>4.1 Fejkilder</b> .....	<b>7</b>
<b>5. DATA</b> .....	<b>9</b>
<b>5.1 Stikprøve-adresser</b> .....	<b>9</b>
<b>5.2 Adresseprojektdata for 4 kommuner</b> .....	<b>9</b>
<b>5.3 Oplysninger om antal adresser og antal lodder på ejerlav, matrikelnr; samt oplysninger om land/by</b> .....	<b>10</b>
<b>5.4 FLAT-adresser</b> .....	<b>10</b>
<b>5.5 Tilknytning af lod id til adresseprojekt- og FLAT adresser</b> .....	<b>11</b>
<b>6. METODE OG RESULTATER</b> .....	<b>11</b>
<b>6.1 Test af målsætninger</b> .....	<b>12</b>
6.1.1 Målsætning 1 Fuldstændighed af adressemængde. ....	12
6.1.1.1 Resultat - Eksistens af adresser. ....	12
6.1.1.2 Resultat - Validitet af adresser .....	13
6.1.2 Målsætning 2 Fuldstændighed af stedfæstelse. ....	13
6.1.2.1 Resultat – Stedfæstelse.....	13
6.1.3 Målsætning 3 Beliggenhed på korrekt matrikelod .....	14
6.1.3.1 Resultat – FLAT’s beliggenhed på korrekt matrikelod.....	14
6.1.4 Målsætning 4 Afstand til korrekt placeret adresse for adresser, der er placeret på korrekt matrikelod.....	17
6.1.4.1 Resultat – Afstand til korrekt placeret adresse .....	17
<b>6.2 Faktorens betydning for forkert placerede adresser</b> .....	<b>17</b>
<b>6.3 Anbefalinger til manuel årsagsforklaring</b> .....	<b>19</b>
6.3.1 Målsætning 1.....	19
6.3.2 Målsætning 3.....	19
<b>7. KONKLUSION</b> .....	<b>20</b>

<b>8. BILAG 1: VÆGTNING AF STIKPRØVE-ADRESSER .....</b>	<b>22</b>
<b>9. BILAG 2: VÆGTNING AF ADRESSEPROJEKTADRESSER.....</b>	<b>23</b>
<b>10. BILAG 3 AFSTAND TIL KORREKT PLACERING .....</b>	<b>24</b>

## 1. Indledning

Foreløbig landsdækkende adressetema (FLAT) er et datasæt der består af ca. 2. millioner adresser stedfæstet omtrentligt ved hjælp af oplysninger fra krydsreferenceregistret (KRR) og det digitale matrikelkort samt i en vis udstrækning Kort & Matrikelstyrelsens topografiske kortdatabase, ”Top10DK”. Stedfæstelsen af de enkelte adresser i FLAT sker ved at udnytte, at en given adresse i KRR er koblet til et bestemt matrikelnummer og at adressen dermed kan stedfæstes til et punkt inden for det pågældende matrikulære areal, hvis geometriske beliggenhed hentes fra det digitale matrikelkort.

Kort & Matrikelstyrelsen (KMS) etablerer FLAT som et led i det såkaldte ”adresseprojekt” der gennemføres af kommunerne i Danmark med bistand fra bl.a. KMS og By- og Boligministeriet. I forbindelse med etableringen af FLAT ønsker KMS at kunne dokumentere kvaliteten af data i en ”varedeklaration”, der beskriver fuldstændighed samt absolut så vel som relativ geometrisk beliggenhed.

Carl Bro as har anbefalet en metode til vurdering af disse data mod et fuldt gennemført adresseprojekt og mod en stikprøve af adresser udvalgt blandt alle landets adresser. Carl Bro har desuden udført selve kontrollen.

Dette dokument er resultatet af undersøgelsen. Dokumentet beskriver de opstillede kvalitetsmål, og indeholder herefter en vurdering af produktionsmetode og datagrundlag. Endelig beskriver dokumentet den anvendte målemetode og resultaterne af kvalitetsmålingen.

## 2. Resume

Foreløbig landsdækkende adressetema (FLAT) er adresser, der er stedfæstet omtrentligt ved hjælp af oplysninger fra krydsreferenceregistret (KRR) og det digitale matrikelkort samt i en vis udstrækning Kort & Matrikelstyrelsens topografiske kortdatabase, ”Top10DK”. Metoden til generering af adresser er udviklet af Kort & Matrikelstyrelsen (KMS). I nærværende rapport er produktionsmetoden evalueret og kvaliteten af de producerede data er beskrevet.

Kvalitetsbeskrivelsen tager udgangspunkt i fire kvalitetsmålsætninger, der er opstillet af KMS. Målsætningerne vedrører fuldstændighed af adresse-mængde, fuldstændighed af stedfæstelse, beliggenhed på korrekt matrikel og afstand til den korrekt placerede adresse for adresser, der er placeret på korrekt matrikel.

Evaluering af produktionsmetoden har vist, at kvaliteten af de producerede adresser må forventes at afhænge af tre hovedfaktorer. *Kvaliteten af rå-data* vurderes særligt at have betydning for kvaliteten af adresser kommuner imellem. Det må således forventes at rå-data, og dermed kvaliteten af de producerede adresser, vil afspejle den indsats, den enkelte kommune har gjort for at ajourføre og verificere de grundlæggende adresse data. *Tiden* vurderes at have indflydelse på kvaliteten i kraft af den tid, der går fra en ændring konstateres i kommunen, til registrene er ajourført; i kraft af den tidsmæssige forskydning mellem udtræk af de enkelte inputdata og i kraft af den tid, der går fra udtræk af baggrundsdata, og til de producerede FLAT adresser anvendes (”henfaldstiden”). I selve *metoden* gennemløber den enkelte adresse en række krav indtil den enten opfylder kravet og dermed stedfæstes eller bliver klassificeres som ikke stedfæstet. Kravene er baseret på antagelser, der giver den største sandsynlighed for korrekt stedfæstelse først i processen. Adresser tilhørende én adresse med én bygning må således forventes at have en højere kvalitet end adresser tilhørende ejendomme med flere matrikelodder og bygninger.

Kvalitetsbeskrivelsen er udført på grundlag af 2 sæt reference adresser. Det ene datasæt er udvalgt som en stikprøve blandt landets adresser. Det andet udgøres af adresser fra fire kommuner, der har

gennemført adresseprojektet. Resultaterne i forhold til de opstillede kvalitetsmål fremgår af nedenstående tabel.

Nr	Mål		Resultat		
	Beskrivelse	Værdi	FLAT internt	Adresseprojekt*	Stikprøve*
1	Andel af reference adresser der kan genfindes i FLAT	> 99%	-	97,6 %	97,6 - 98,6 %
	Andel af FLAT der ikke kan genfindes i referenceadresser	< 1%	-	3,7 – 3,9 %	-
2	Andel af FLAT der er stedfæstet	> 98%	96,3 %	-	-
3	Andel stedfæstede FLAT på korrekt matrikelod	> 99,8	-	94,3 – 94,6 %	93,6 - 95,3 %
	Andel FLAT på forkert matrikelod med afstand til korrekt placering på max 200 m	Hovedparten	-	76,3 – 79,5 %	72,3 - 86,7 %
4	Andel af by FLAT på korrekt matrikelod med maks afstand på 100 m til korrekt beliggenhed på lod	Hovedparten	-	94,1 – 94,4 %	-
	Andel af land FLAT på korrekt matrikelod med maks afstand på 250 m til korrekt beliggenhed på lod	Hovedparten	-	98,1 – 98,6 %	-

\* Værdierne er vægtede, således at det er de forventede værdier for hele landet, der er angivet i tabellen

Resultaterne viste sig for nogle måls vedkommende væsentligt bedre end de opstillede forventninger, mens de for andre viste sig knap så gode. Et udsnit af de adresser, der ikke var stedfæstet korrekt bliver af KMS gennemgået manuelt med henblik på at finde årsager og systematiske fejl og herigennem opnå mulighed for at forbedre metoden.

### 3. Kvalitetsbeskrivelsens grundlag

Hensigten med kvalitetsbeskrivelsen er at give en statistisk beskrivelse af hvorledes FLAT adresser lever op til de opstillede kvalitetsmålsætninger, således det er muligt at vurdere FLAT-adressernes anvendelighed og metodens brugbarhed i forhold til en konkret opgave. I denne forbindelse lægger KMS især vægt på at kunne afdække og kvantificere de fejlkilder som undersøgelsen afslører.

Udarbejdelse af kvalitetsbeskrivelsen for FLAT-adresser har taget udgangspunkt i følgende:

1. KMS's notat af 21. september 1999 "Målsætning for FLAT-varedeklaration"
2. KMS's principdiagram fra september 1999 for FLAT adressers produktionsmetode
3. Møder mellem KMS og Carl Bro den 14. og 28. okt. 1999 samt 26. jan. og 11. feb. 2000

#### 3.1 Kvalitetsmålsætning for FLAT adresser

Ideelt set burde det referencegrundlag som FLAT-adresserne måles i forhold til, være den faktisk fysiske "adressevirkelighed". I praksis er dette naturligvis ikke muligt. Som den bedste alternative løsning er der i stedet, som beskrevet i afsnit 5, udpeget et sæt kommunale reference-adresser der kan forventes at have den højest tilgængelige kvalitet.

Der er opstillet fire kvalitetsmålsætninger for FLAT adresserne:

##### 3.1.1 Målsætning 1: Fuldstændighed af adressemængde

Målsætning 1 er udtryk for hvor fuldstændige FLAT adresserne er i forhold til de adresser, der findes i et originalt JF50 udtræk for kommuner, der har gennemført adresseprojektet. Målsætningen giver både

et udtryk for FLAT's andel af eksisterende adresser, men også et udtryk for i hvor høj grad det kan forventes at FLAT adresser rent faktisk eksisterer (at FLAT ikke har et højt indhold af "administrative adresser").

- *Mindst 99% af adresserne i referencegrundlaget skal kunne genfindes i FLAT adresserne.*
- *Højest 1 % af adresserne i FLAT kan ikke genfindes i referencegrundlaget.*
- *Hovedparten af uoverensstemmelser mellem FLAT og referenceadresser kan forklares ved forskel i udtrækstidspunkt og deciderede fejl i de data der ligger til grund for FLAT. Dvs. fejl, der skyldes andet end produktionsmetoden.*

### 3.1.2 Målsætning 2: Fuldstændighed af stedfæstelse

Målsætning 2 er udtryk for den andel af FLAT adresser, der er stedfæstet ved et koordinatsæt.

- *Mindst 98% af FLAT adresserne er stedfæstet ved beregnede koordinater*

### 3.1.3 Målsætning 3: Beliggenhed på korrekt matrikel lod

Målsætning 3 er udtryk for den andel af FLAT adresser, der ved stedfæstelse på det digitale matrikelkort bliver placeret i det matrikel lod hvori den tilsvarende referenceadresse placeres. En adresse som i BBR refererer til en bestemt bygning eller opgang i en bygning skal således placeres på samme matrikel lod som bygningen fysisk er beliggende på. Hvis en adresse alene refererer til en ubebygget ejendom, skal den tilsvarende være placeret på den pågældende fysiske ejendom. Hvis en ubebygget ejendom består af flere matrikel lodder, bør adressen være placeret på den lod, der ligger nærmest naboadressen.

- *Mindst 99,8% af stedfæstede FLAT-adresser ligger på korrekt matrikel lod.*
- *For adresser på forkert matrikel lod har hovedparten af FLAT-adresserne en afstand til den korrekte beliggenhed på max. 200 m.*
- *Hovedparten af de forkert placerede adresser kan forklares ved forskel i udtrækstidspunkt og deciderede fejl i de data der ligger til grund for FLAT.*

### 3.1.4 Målsætning 4: Afstand til den korrekt placerede adresse for adresser, der er placeret på korrekt matrikel lod

Målsætning 4 er udtryk for den gennemsnitlige afstand fra en FLAT adresse til adressens korrekte beliggenhed i den rigtige bygning i overensstemmelse med TK-standarden. Det følger af FLAT-produktionstidspunktmetoden, at denne eksakt rigtige beliggenhed ikke kan opnås, idet FLAT-adressen som udgangspunkt placeres i den største bygning på det valgte matrikel lod. Ved ejendomme med flere bygninger, f.eks. etagebolig- og andelsboligbebyggelser, erhvervs- og institutionsbyggeri samt større landbrugsejendomme, kan det medføre at adressen placeres "langt" fra den korrekte beliggenhed.

- *Afstanden fra en FLAT-adresse på korrekt matrikel lod til den korrekte beliggenhed er for hovedparten af by adresser mindre end 100 meter og for land adresser mindre end 250 meter.*

## 4. Vurdering af produktionsmetoden i forhold til målsætningerne

Udgangspunktet for FLAT-metoden, er at der dannes et såkaldt JF50-udtræk på basis af Krydsreferenceregisteret (KRR), og et "forbedret Geonøgle M" udtræk på basis af det digitale matrikelkort og bygningstemaet fra TOP10DK.

JF50-udtrækket rummer een record for hver adresse og angiver samtidig hvilket matrikelnummer adressen hører til på og om adressen hører til en bygning eller en bolig- eller erhvervsenhed i BBR.

”Forbedret Geonøgle M” er et datasæt med een record pr lod i matrikelkortet. Ud over LodId indeholder datasættet, ejerlavs- og matrikelbetegnelse samt koordinaterne til et punkt, der ligger inden for loddens flade. Hvis der ligger en eller flere TOP10DK bygninger eller dele af bygninger på lodden (>25 m<sup>2</sup>), er dette punkt flyttet ind i den største af disse bygninger.

I produktionen af FLAT gennemgås en række aktiviteter. Afhængig af hvordan en adresse opfylder den enkelte aktivitet kan adressen enten fortsætte i processen, tilknyttes et koordinatsæt med kvalitetsbetegnelse eller blive klassificeret som ikke stedfæstet.

Hovedprincipperne i processen er:

1. Hvis KRR adressen findes i Geonøgle M påsættes koordinater herfra på adressen fra KRR Ellers klassificeres som ikke stedfæstet.
2. Hvis der er præcis et lod til det matrikelnummer, der hører til adressen placeres adressen på dette lod og klassificeres som M1. Ellers fortsættes til 3.
3. Hvis der er et eller flere lodder, der er bebygget, placeres adressen på lod med største bygning og klassificeres som M2. Ellers fortsættes til 4.
4. Adressen placeres på det lod, der er tættest ved nabo og klassificeres som M3.

Da denne proces i nogle tilfælde giver åbenlyse fejl, indeholder metoden tre metoder til supplerende kontrol af hovedprincipperne:

1. Adresser, der i KRR peger på et forkert matrikelnummer.  
M1 adresser, der iflg. BBR er bebygget, men som iflg. TOP10DK ikke ligger på et bebygget lod, udvælges. Hvis der på de andre matrikelnumre, der iflg. KRR hører til ejendommen, findes TOP10DK bygninger placeres adressen på matrikelnummer/lod med den største bygning og klassificeres som M4, Adressen flyttes dog maks. 1 km. Ellers bevares M1 klassifikationen.
2. Adresser der er placeret på største bygning, men hvor adressen tydeligvis ligger for langt væk fra naboadresserne. Kriteriet ”tydeligvis for langt fra naboer” vurderes ud fra om vurderingsejendommen har andre matrikler, der ligger tættere på naboen end den hvorpå adressen er blevet placeret.  
M2 og M4 adresser, der er bebygget iflg. BBR udvælges. Hvis der ikke på de andre matrikelnumre, der iflg. KRR hører til ejendommen, findes TOP10DK bygninger, bevares den oprindelige klassifikation. Ellers vurderes om en eller flere af kandidaterne ligger tættere på adressens naboer end den valgte adresse. Hvis dette er tilfældet klassificeres adressen som ikke stedfæstet. Ellers bevares den oprindelige klassifikation.
3. Adresser der ligger ”usandsynligt” langt fra naboer.  
Hvis afstanden fra en adresse til de to naboer er over 2,5 km klassificeres adressen som ikke stedfæstet.

FLAT adresser produceres ikke for Københavns- og Frederiksberg kommuner.

#### 4.1 Fejlkilder

FLAT adresserne er et produkt af de data, der bruges til sammenstillingen og produktions metoden. Kvaliteten af FLAT er derfor også afhængig af disse faktorer.

Fejlkilder i produktionen af FLAT kan opdeles i tre hovedgrupper:

1. Kvaliteten af de data, FLAT-adresser dannes ud fra
2. Tiden
3. Metoden, der anvendes til dannelse af FLAT-adresser

**ad 1)** Basisdata for FLAT er først og fremmest adresseoplysningerne fra krydsreferencesystemet, der som et såkaldt JF50-udtræk.

Det centrale led i dannelsen af JF50-udtrækket er finde frem til hver addresses matrikelnummer ved hjælp af oplysningerne i Bygnings- og Boligregisteret (BBR) og krydsreferenceregisteret (KRR). Metoden indebærer, at adresser der befinder sig på en ejendom bestående af eet matrikelnummer, placeres på dette. På ejendomme bestående af flere matrikelnumre, anvendes først BBR-oplysningen om hvilken bygning, adressen hører til og dernæst KRR-oplysningen om hvilket matrikelnummer denne bygning er placeret på. Især den sidste reference mellem bygning og matrikelnummer udgør en fejlkilde, idet der er tale om en reference der blev dannet maskinelt i forbindelse med etableringen af KRR i 1989, og som kommunerne ikke i alle tilfælde har berigtiget.

Med hensyn til adresseoplysningernes fuldstændighed i KRR - og dermed i JF50-udtrækket - kan man ikke forvente at kvaliteten er ensartet kommunerne imellem. Forskellene vil bl.a. afhænge af hvor komplekse adresseforholdene er i de enkelte kommuner (byområder er mere komplekse end landområder), og hvorledes arbejdet med ajourføring af adresser i ESR, BBR og KRR er tilrettelagt i kommuner. Det må antages, at KRR er mere korrekt i kommuner, der har gennemført adresseprojektet end i kommuner, der ikke har.

Tilsvarende kan der være forskel på kvaliteten af de enkelte "adresstyper"; f.eks. er der erfaring for, at der er flere administrative adresser og lignende "snavs" i de adresser der alene stammer fra ESR (dvs. adresser til ubebyggede ejendomme) end i de bebyggede adresser der findes i BBR.

KRR kan antages at være næsten fuldstændigt og korrekt med hensyn til matrikelidentifikationer. I KRR anvendes dog 3-cifrede, "kommunale ejerlavskoder", mens der i Geonøgle M anvendes 6-7 cifrede "landsejerlavskoder". Indtastningen af kommunale ejerlavskoder i ESR sker ved en manuel proces, der kan være behæftet med fejl.

Forbedret Geonøgle M antages at være korrekt og uden regionale variationer eller variationer mellem forskellige områdetyper (fx. by og land).

**ad 2)** Tiden har på flere niveauer betydning for kvaliteten af de producerede FLAT adresser:

- Den tid der går fra en ændring konstateres i kommunen til registeret er ajourført.
- Den tidsmæssige forskydning mellem udtræk af de enkelte inputdata til FLAT dannelsen.
- Den tid der går fra udtræk af baggrundsdata og til de producerede FLAT adresser anvendes – "henfaldstiden".

**ad 3)** FLAT-metoden placerer altid adresserne i den største bygning på lodden. I områder hvor forholdet mellem TOP10DK-bygninger og adresser på en lod er 1:1 eller hvor "den rigtige bygning" samtidig er den største, vil denne metode i de fleste tilfælde give et godt resultat. Dette vil typisk være tilfældet i områder, der er domineret af enfamiliehuse, rækkehuse, karrébebyggelse og sommerhuse.

I situationer hvor der derimod findes to eller flere bygninger pr. lod, vil det nødvendigvis være helt tilfældigt, om adressen kommer til at sidde på den rigtige bygning. Hvis der f.eks. er tale om en landbrugsejendom, vil metoden som regel placere adressen på en af avlsbygningerne, idet stuehuset oftest er den mindste eller den næstmindste bygning.

Tilsvarende vil situationer hvor mere end en adresse hører til på en og samme lod, resultere i at alle adresser placeres med udgangspunkt i samme bygningscentroide. I de to nævnte situationer vil FLAT-metoden altså indebære at en adresse sidder inden for den rigtige lod, men ikke på den rigtige bygning.

Fejlmulighederne øges imidlertid i situationer hvor et matrikelnummer består af flere lodder. FLAT-metoden placerer pr. definition adressen på den lod, der rummer den største bygning (M2). Denne regel

må nødvendigvis give forkerte resultater i nogle tilfælde – specielt for landbrugs- og industriejendomme. Alternativet hertil er at placere adressen på lod tættest ved nabo (M3).

Betydningen af bl.a. denne fejlkilde søges reduceret gennem et princip om at en adresse hellere må undvære en stedfæstelse end at have en forkert stedfæstelse. Vurderingskriterierne i de tre supplerende kontrolmetoder er således sat ud fra et ønske om at fjerne stedfæstelsen på adresser, hvis stedfæstelse der kan være tvivl om. Dette kan have en negativ betydning for fuldstændigheden af stedfæstede adresser, men omvendt en positiv indflydelse på de adresser, der er stedfæstet. Det må forventes at tvivlstilfældene oftes kan henføres til adresser, der hører til ejendomme med flere matrikler og ejendomme hvor den adressebærende bygning ikke er den største. Betydningen af de supplerende kontrolmetoder vil derfor være størst for denne type adresser.

## 5. Data

Der indgår to referencedatasæt i undersøgelsen. Det ene består af en stikprøve af adresser spredt over det meste af landet (herefter kaldet stikprøve-adresser), mens det andet består af et fuldt gennemført adresseprojekt i fire kommuner.

### 5.1 Stikprøve-adresser

Stikprøve-adresser er dannet til et andet projekt for en statslig myndighed i 1997. Grundlaget for stikprøve-adresser er en udvælgelse af 10-20 adresser i samtlige kommuner uafhængigt af kommunens størrelse. Ejere af de pågældende ejendomme er spurgt, om de ønskede at medvirke i en given undersøgelse. Adresser, for de der ikke ønskede dette, blev fjernet. Kommunerne har herefter manuelt indhentet oplysninger om beliggenhed. Afslutningsvis er adresserne digitaliseret på grundlag af indtegnning på 4-cm kort eller teknisk kort.

Da stikprøve-adresserne pr. definition er bebyggede, kan de ikke anvendes til test af om bebygget/ubebygget har betydning for FLAT adressernes kvalitet. Dette betyder endvidere, at konklusioner fra test med anvendelse af stikprøve-adresser kun kan drages for bebyggede adresser.

Datasættet indeholder bl.a. oplysninger om:

"komnr";  
"vejkode";  
"husnrbgst";  
"x-koordinat";  
"y-koordinat";  
"ELAVSKODE";  
"ELAVNAVN";  
"MAS\_MATRNR";  
"LOD\_ID"

Stedfæstelsen af stikprøve-adresser er foretaget i System 34. I alt anvendes 2943 stikprøve-adresser som reference.

### 5.2 Adresseprojektdata for 4 kommuner

I testen anvendes data fra 4 kommuner, der har gennemført adresseprojektet: Odense, Helsingør, Brande og Hjørring. De 4 kommuner er udvalgt ud fra en forventning om, at de repræsenterer kommuner der har udført et omhyggeligt arbejde med adresseprojektet. Kommunerne repræsenterer samtidig forskellige kommunetyper: Store "Adressekommuner", provinskommuner og landkommuner.

I de modtagne adresseprojektdresser er der fundet enkelte dobbeltforekomster, dvs. adresser for hvilke der er mere end én stedfæstelse. Disse er ligeledes udeladt af den videre undersøgelse.

Der forefindes oplysninger om:

"Kommunenr";  
"Ejerlav";  
"Matrikelnr.";   
"Vejkode";  
"Husnrbogstav";  
"Vejnavn";  
"X";  
"Y";  
"Husnr".

Ialt anvendes 108458 adresser fra adresseprojektet.

### 5.3 Oplysninger om antal adresser og antal lodder på ejerlav, matrikelnr; samt oplysninger om land/by

På grundlag af matrikelkortet er der, til brug for årsagsforklaring, etableret et datasæt med karakteristika for hver matrikel. Data indeholder oplysninger om antal lodder til en matrikel, antal adresser til en matrikel og beliggenhed i land/by. Oplysningerne om hvorvidt en matrikel er beliggende i land eller by er generet på grundlag af bypolygoner i digitalt oversigtskort (1:200.000)

Der forefindes oplysninger om:

"Ejerlav";  
"Matbetegnelse";  
"Antal adresser";  
"Antal lodder";  
"Land/by/både land og by".

### 5.4 FLAT-adresser

FLAT indeholder bl.a. to felter ("antbyg" og "antenh"). Når begge disse felter for en adresse er 0 eller tomme, ligger adressen på et ubebygget lod.

Der forefindes oplysninger om:

"Kommunenr";  
"Vejkode";  
"Vejnavn";  
"Husnrbogstav";  
"Postnr";  
"Postdistrikt";  
"Ejerlav";  
"Matrikelnummer";  
"Revdato";  
"Y-koordinat";  
"X-koordinat";  
"husnr";  
"antbyg";  
"antenh";  
"N-klasse"

Foruden København- og Frederiksberg Kommuner har det ikke været muligt at inddrage Århus og Randers Kommuner i undersøgelsen, da FLAT ikke var produceret for disse på undersøgelsestidspunktet.

Manglen af FLAT for Århus og Randers kommuner har, i forhold til undersøgelsen, betydning i den udstrækning adresserne fra disse to kommuner adskiller sig fra gennemsnittet. Da både Randers og Århus Kommuner må antages at have en større andel byadresser end gennemsnitligt for landet vil

resultaterne ikke umiddelbart kunne ekstrapoleres til hele landet. Af praktiske hensyn er det alligevel valgt at antage, at fordelingen af adressetyper i FLAT-adresser er i overensstemmelse med fordelingen af landets adresse typer. Vægtningen af adresser påvirkes derfor både i forhold til stikprøve-adresser og adresseprojektadresser, da landsfordelingen beregnet på grundlag af FLAT anvendes til vægtning.

### 5.5 Tilknytning af lod id til adresseprojekt- og FLAT adresser

For at kunne teste målsætning 3 er adresseprojekt adresser og FLAT adresser påhæftet oplysning om det lod i hvilket adressen er stedfæstet. For FLAT adresser er det dog kun de adresser, der matcher stikprøve-adresser eller adresser i de 4 testkommuner, der er påhæftet oplysningen. Oplysningen er påsat ved at sammenstille adressernes stedfæstelse med oplysninger fra det digitale matrikelkort gennem en geografisk analyse.

## 6. Metode og resultater

Udgangspunktet for udarbejdelse af en metode til kvalitetsbeskrivelse af FLAT adresser har været en evaluering af hvilke faktorer, der vurderes at have indflydelse på hvorvidt kvalitetsmålsætningerne kan opfyldes. De væsentligste faktorer, der er blevet lagt til grund for metoden er:

- Antallet af lodder pr. matrikel, således at andelen af adresser, der er placeret på korrekt matrikel, forventes at falde med stigende antal lodder (industri og landbrug har typisk flere lodder pr. matrikel).
- Antallet af adresser pr. matrikel, således at der forventes en faldende absolut nøjagtighed i stedfæstelsen med stigende antal adresser pr. matrikel.
- Bygningsstrukturen, således at der forventes en faldende relativ og absolut nøjagtighed for adresser, hvor den adressebærende bygning ikke er den største (typisk landbrug og industri).
- Kommunens aktiviteter og rutiner omkring adressevedligeholdelse, således at kvaliteten af FLAT forventes at stige med stigende aktivitet omkring adressevedligeholdelse.
- Størrelsen af kommune, idet der forventes større kompleksitet i adresserne for de gamle bykerner i større kommuner og dermed større usikkerhed for adresserne i disse områder.

Udover at referencedata optimalt skal kunne give svar på betydningen af de listede faktorer har repræsentativiteten af referencedata været overvejet. Spørgsmålene er bl.a.:

- Er KRR bedre for de kommuner, der har gennemført adresseprojektet?
- Kan resultater fra kvalitetsmålinger med adresseprojekt adresser som reference overføres til kommuner, hvor adresseprojektet ikke er gennemført?
- Hvad er fuldstændigheden for referencegrundlaget?
- Er fuldstændigheden for referencegrundlaget ens kommunerne imellem?
- Er fuldstændigheden for referencegrundlaget ens adressetyper imellem?
- Er der forskel i udtrækstidspunkt for reference henholdsvis FLAT-adresser?
- Kan referencegrundlaget betragtes som ajourført på udtrækstidspunktet?
- Hvor fuldstændigt er referencegrundlaget med hensyn til stedfæstelse?
- Er fuldstændigheden mht. stedfæstelse for referencegrundlaget ens kommunerne imellem?
- Er fuldstændigheden mht. stedfæstelse for referencegrundlaget ens adressetyper imellem?
- Hvor fuldstændigt er referencegrundlaget med hensyn til placering på korrekt matrikel?
- Er graden af korrekt placering for referencegrundlaget ens kommunerne imellem?
- Er graden af korrekt placering for referencegrundlaget ens adressetyper imellem?
- Hvilken betydning har den tidsmæssige forskydning mellem udtræk af referencegrundlag og af baggrundsdata?

Der har været overvejet flere modeller for etablering af et repræsentativt reference grundlag. Fælles for modellerne var, at de forudsatte en høj grad af validering af reference adresserne i forhold til den fysiske virkelighed. Da ajourførte adresseprojekt adresser i dag kan betragtes som de bedst tilgængelige

stedfæstede adresser, blev det vurderet, at de overvejede modeller ikke gav en merkvalitet, der stod mål med de omkostninger, der ville være forbundet med etablering af et særligt referencedatasæt og med den tidsmæssige betydning. Det blev derfor besluttet at anvende adresse projektdata fra de 4 kommuner nævnt i afsnit 5.2. Derudover blev det besluttet, at supplere med stikprøve adresser nævnt i afsnit 5.1. Da disse adresser er en stikprøve udvalgt blandt samtlige kommuner og da stedfæstelsen er sket manuelt for den enkelte adresse, er disse et godt supplement til adresseprojekt adresserne.

Stikprøve-adresserne er altså produceret uafhængigt af FLAT og adresseprojekt adresserne både med hensyn til tidspunkt og metode. I de test hvor begge typer reference anvendes er der derfor tale supplerende kvalitetsbeskrivelser, der ikke umiddelbart kan sammenstilles.

## 6.1 Test af målsætninger

For at sikre den bedst mulige repræsentativitet af testresultaterne er der foretaget en vægtning i de enkelte test ud fra forholdet mellem fordelingen af de enkelte faktorer i referencen i forhold til fordelingen af de tilsvarende faktorer på landsplan, se bilag 1 og 2.

Denne landsdækkende fordeling er foretaget på grundlag af adresser i FLAT og data om antal lodder og adresser pr matrikel, forholdet mellem bebyggede/ubebyggede adresser og adressens beliggenhed i land/by.

Det er vurderet, at FLAT kan anvendes her, idet:

- test af målsætning 1 viste en tilfredsstillende overensstemmelse
- da fordelingen ikke skal anvendes til at se på den enkelte adresse
- da differencen til et "eksakt" antal antages at være beskeden i forhold til behovet (en vægtning)
- da de anvendte bypolygoner antages at være en langt mere betydende fejlkilde.

Se i øvrigt afsnit 5.4 om FLAT's repræsentativitet.

Der anvendes et sikkerhedsinterval på 95%. Alle efterfølgende test af målsætningerne opgøres med et signifikansniveau på 95%, det vil sige testsandsynligheder større end 0,05.

### 6.1.1 Målsætning 1 Fuldstændighed af adressemængde.

Der er foretaget en test af FLAT adresser mod adresseprojekt adresser og en test mod stikprøve-adresser.

1. Beregning af den andel af adresserne fra adresseprojekt og stikprøve-adresser der findes i FLAT adresserne. Blandt den del af reference adresserne der ikke kan genfindes i FLAT adresserne er udvalgt et antal. Årsagen til uoverensstemmelse klarlægges af KMS ved manuel gennemgang.
2. Beregning af den andel af adresserne fra FLAT der ikke findes i adresseprojekt adresserne. Blandt de FLAT adresser, der ikke kan genfindes i reference adresserne er udvalgt et antal. Årsagen til uoverensstemmelse er klarlægges af KMS ved manuel gennemgang.

Deltest 2 er ikke udført i forhold til stikprøve-adresser, da disse netop kun udgør en stikprøve.

#### 6.1.1.1 Resultat - Eksistens af adresser.

Der er undersøgt med hvor stor sikkerhed man kan genfinde en reference adresse i FLAT. En adresse er her defineret ved kommunenr, vejkode og husnr med evt efterfølgende bogstav. Det er derfor underordnet om adresserne er stavet ens.

#### Stikprøveadresser

Til undersøgelse af eksistens af adresser på landsplan, er der benyttet den landsdækkende stikprøve på 2943 adresser. Disse adresser er vægтет på baggrund af fordelingen på adrestyper for hele landet.

37 adresser af de 2943 stikprøveadresser kunne ikke findes blandt FLAT-adresserne. Dette svarer til, at 37879 adresser ud af 1989103 adresser på landsplan ikke vil kunne findes i FLAT.

Med 95% sikkerhedsintervaller vil man kunne forvente, at mellem 1,4% og 2,4% af adresserne ikke vil kunne genfindes i FLAT. Dette svarer til at:

- *Mellem 97,6 og 98,6 % af stikprøve-adresserne kan genfindes i FLAT adresserne.*

### **Adresseprojektadresser**

Eksistensen er også undersøgt mod adresseprojektadresser for de 4 kommuner, hvor der tilsammen er 108458 adresseprojekt adresser. Dette giver ved vægtning på baggrund af fordelingen på adresstyper for hele landet (bilag 2) en adresse mængde på 1989097.

626 adresser af de 108458 kunne ikke findes blandt FLAT-adresserne. Dette svarer til at 47614 adresser på landsplan ikke vil kunne findes i FLAT.

Med 95% sikkerhedsintervaller vil man kunne forvente, at mellem 2,38% og 2,40% af adresserne ikke vil kunne genfindes i FLAT. Dette svarer til at:

- *97,6 % af adresseprojektadresserne kan genfindes i FLAT adresserne.*

Den store forskel på vægtede (97,6) og uvægtede (94,2%) data skyldes primært, at der i de fire kommuner er en meget stor overrepræsentativitet af byadresser.

#### **6.1.1.2 Resultat - Validitet af adresser**

Det er ikke nok, at vide med hvor stor sikkerhed man vil kunne genfinde en adresse i FLAT. Hvis man vil anvende data i FLAT vil man i mange tilfælde også være interesseret i at vide, i hvor stor udstrækning datamaterialet i FLAT repræsenterer adresser, der rent faktisk eksisterer.

Der er undersøgt, hvor stor en del af FLAT-adresser beliggende i de fire Adresseprojektkommuner, der kunne genfindes i Adresseprojektadresserne. I FLAT var ialt 109503 adresser beliggende i de fire kommuner.

1671 adresser af de 109503 adresser kunne ikke genfindes blandt Adresseprojektadresserne. Dette svarer ved vægtning på landsplan til at 75669 ud af 1989097 adresser i FLAT ikke vil kunne findes i Adresseprojektadresserne.

Den store forskel mellem vægtet og uvægtet skyldes, at fordelingen af adresseprojektadresser er skæv i forhold til FLAT adresserne (bilag 2). Resultatet tyder således på at en del af de 1671 adresser tilhører en type der er underrepræsenteret i Adresseprojektadresserne (tildeles høj vægt).

Med 95% sikkerhedsintervaller vil man kunne forvente, at:

- *Mellem 3,7% og 3,9% af adresserne i FLAT ikke kan genfindes i adresseprojektadresserne.*

#### **6.1.2 Målsætning 2 Fuldstændighed af stedfæstelse.**

Det er beregnet hvor stor en del af FLAT adresserne, der er stedfæstet ved et koordinatsæt.

##### **6.1.2.1 Resultat – Stedfæstelse**

72703 FLAT adresser svarende til 3,7 % manglede koordinater i FLAT. Dette svarer til at:

- *96,3 af FLAT adresserne er stedfæstet ved beregnede koordinater*

### 6.1.3 Målsætning 3 Beliggenhed på korrekt matrikellod

For hver adresse er lavet en sammenstilling af adresseprojekt adresser og stikprøve-adresser mod FLAT adresser. På grundlag heraf er beregnet, hvor stor en del af de stedfæstede FLAT adresser, der har samme LodID i referencen og i FLAT.

For de FLAT adresser, der ikke er korrekt stedfæstet (identiske LodID) er afstanden til den korrekte placering beregnet gennem en sammenstilling af FLAT adresser og referenceadresser ved adresseID. Afstanden er beregnet som fugleflugtlinie.

Blandt de forkert placerede adresser er udtrukket et antal til manuel kontrol af årsagen til forkert placering.

#### 6.1.3.1 Resultat – FLAT's beliggenhed på korrekt matrikellod

##### Stikprøveadresser

123 af de 2906 (2943-37) stikprøveadresser, der kan genfindes i FLAT, havde ikke samme lod-id som den tilsvarende FLAT-adresse. Dette svarer på landsplan til, at 106793 adresser ud af 1922420 stedfæstede adresser er beliggende på korrekt matrikellod.

Med 95% sikkerhedsintervaller vil man kunne forvente, at mellem 4,72% og 6,40% af de stedfæstede adresser ikke vil være placeret på korrekt matrikellod.

Afstanden fra FLAT til stikprøveadresserne er beregnet for de 123 adresser, der var placeret på forkert lod. Fordelingen af de vægtede data er vist i Figur 1 og 2.

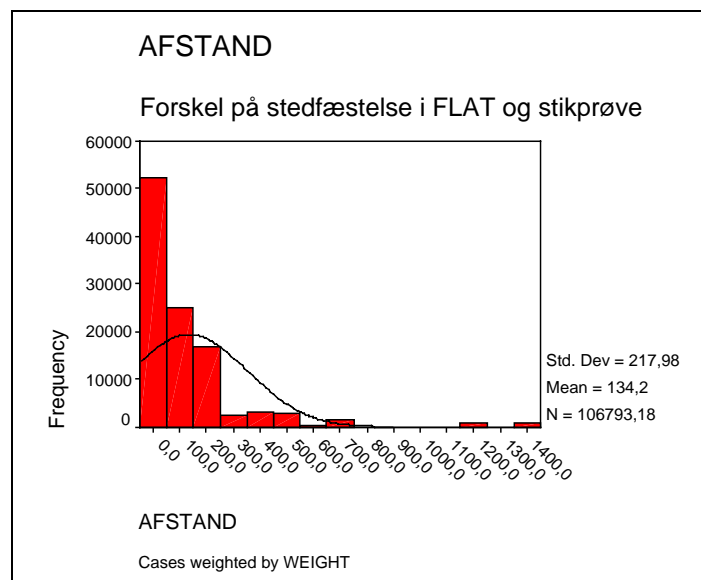


Figure 1

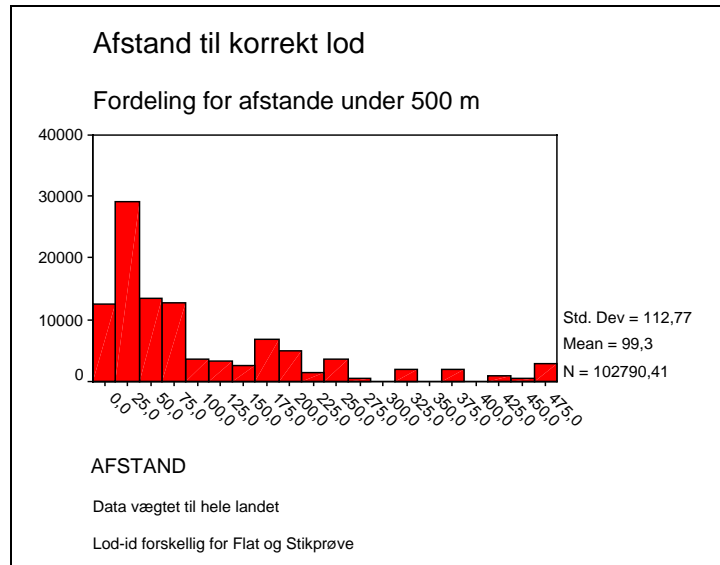


Figure 2

Den gennemsnitlige afstand er 134,2 m med en standardafvigelse på 217,98 m. 101 af de 123 adresser har en afstand på under 200 m til den korrekte adresse. Dette svarer på landsplan til, at 84865 adresser ud af 106793 adresser er beliggende under 200 m fra stikprøveadressen.

Med 95% sikkerhedsintervaller vil man kunne forvente, at mellem 72,33% og 86,61% af de forkert placerede adresser vil have en afstand på under 200 m fra de korrekt placerede adresser.

Dette svarer til at:

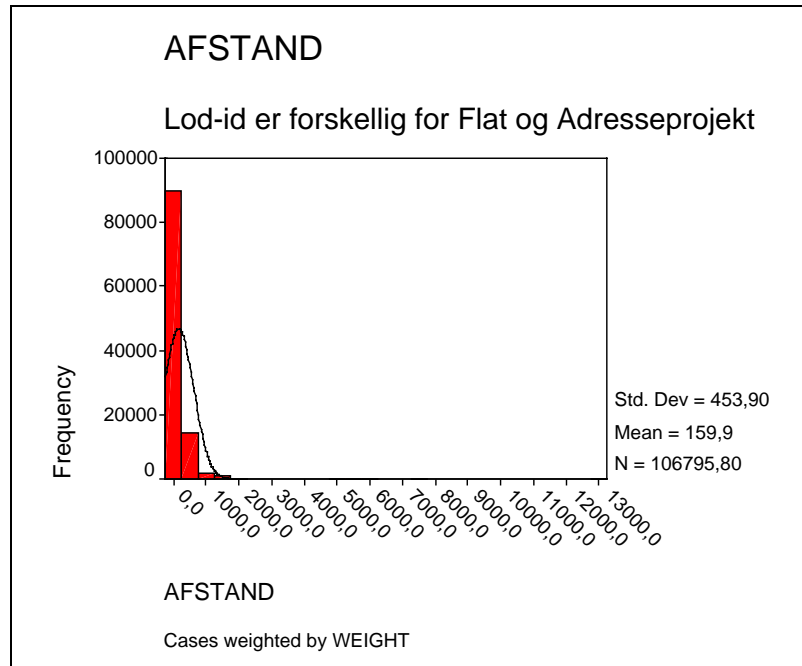
- Mellem 93,6 og 95,3% af de stedfæstede FLAT-adresser ligger på korrekt matrikel lod.
- For forkert placerede adresser har mellem 72,3% og 86,7% af FLAT-adresserne en afstand til den korrekte beliggenhed på max. 200 m.

### Adresseprojektadresser

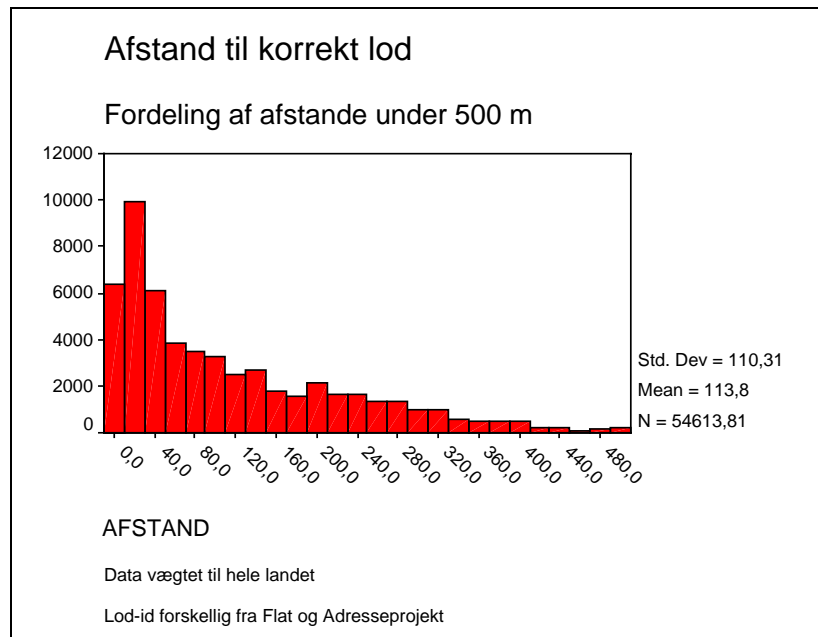
3627 af de 107832 adresser, der både findes i Adresseprojekt adresserne og i FLAT adresserne havde ikke overensstemmende lod-id. Dette svarer på landsplan til, at 106798 adresser ud af 1918850 stedfæstede adresser på landsplan er beliggende på korrekt matrikel lod.

Med 95% sikkerhedsintervaller vil man kunne forvente, at mellem 5,42% og 5,70% af de stedfæstede adresser ikke vil være placeret på korrekt matrikel lod.

Afstanden fra adresseprojektadresserne til FLAT er beregnet for de 3627 adresser, der ikke er beliggende på korrekt lod-id. Fordelingen af de vægtede data er vist i figur 3 og 4.



**Figure 3 Afstanden til den korrekte beliggenhed.  
Den største observation er 13246m**



**Figure 4**

Den gennemsnitlige afstand er 159,9 m med en standardafvigelse på 453,9 m 2675 af de 3627 adresser var beliggende med en afstand på under 200 m til den korrekte adresse. Dette svarer på landsplan til, at 83204 adresser ud af 106798 adresser er beliggende under 200 m fra adresseprojektadressen.

Med 95% sikkerhedsintervaller vil man kunne forvente, at mellem 76,34 % og 79,48 % af de forkert placerede adresser vil have en afstand på under 200 m fra de korrekt placerede adresser.

Dette svarer til at:

- *Mellem 94,3% og 94,6 % af de stedfæstede FLAT-adresser ligger på korrekt matrikel lod.*
- *For forkert placerede adresser har mellem 76,3 % og 79,5 % af FLAT-adresserne en afstand til den korrekte beliggenhed på max. 200 m.*

#### **6.1.4 Målsætning 4 Afstand til korrekt placeret adresse for adresser, der er placeret på korrekt matrikelod**

For adresser, der er fundet placeret på korrekt lod under kvalitetsmål 3 er afstanden til den korrekte beliggenhed på lodden beregnet gennem en sammenstilling af FLAT adresser og adresseprojekt adresser ved adresseID. Afstanden er beregnet som fugleflugtlinie.

##### **6.1.4.1 Resultat – Afstand til korrekt placeret adresse**

102074 adresser blandt de 103873 (107832-3627-332 (forskellig lod id)) adresser, der er beliggende på korrekt lod-id har en afstand til adresseprojektadresserne på mindre end 200 m. 91120 adresser ligger endda indenfor en afstand af 50 m. Dette svarer på landsplan til, at 1787574 ud af 1812052 adresser, der er placeret på korrekt lod, ligger indenfor en afstand af 200 m fra adresseprojektsadressens placering.

Med 95% sikkerhedsintervaller vil man kunne forvente, at:

- *Afstanden fra en FLAT-adresse, der er beliggende på korrekt lod, til den korrekte beliggenhed på lodden er for mellem 98,6 % og 98,7 % af adresserne maksimalt 200 meter.*

Ud fra en forventning om, at afstandene afviger mere på landet end i byen er der sat kvalitetsmål op separat for land- og byadresser.

Der er 90942 byadresser, hvoraf de 90896 er stedfæstede. Af disse er 84215 beliggende indenfor en afstand af 100 m fra adresseprojektadressen. Dette svarer på landsplan til, at 1238645 ud af 1314806 stedfæstede adresser i byen er beliggende indenfor en afstand af 100 m fra adresseprojektplaceringen.

Der er 15457 stedfæstede landadresser, hvoraf 15155 befinder sig indenfor en afstand af 250 m fra placeringen af adressen i adresseprojektet. Dette svarer på landsplan til at 502068 ud af 510697 atedfæstede adresser på landet er beliggende inden for en afstand af 250 m fra placeringen i adresseprojektet. Se i øvrigt bilag 3.

Med 95% sikkerhedsintervaller vil man kunne forvente, at:

- *Afstanden fra en FLAT-adresse på korrekt matrikel lod til den korrekte beliggenhed for mellem 94,1 % og 94,4 % af by adresser er mindre end 100 meter og for mellem 98,1 % og 98,6 % af land adresser er mindre end 250 meter.*

## **6.2 Faktors betydning for forkert placerede adresser**

For referenceadresserne er der undersøgt, hvorvidt faktorer som antallet af adresser på matriklen, antal lodder på matriklen og adressens beliggenhed i by og på land har nogen betydning for, om adressen er blevet placeret på korrekt lod. For adresseprojektadresserne er der tillige undersøgt om det har betydning, om adressen er beliggende på en bebygget matrikel eller ej.

For adresser, der ligger på matrikler med flere adresser er der 84,9 % blandt stikprøveadresserne, der er placeret korrekt, mens det gælder for hele 97,0 % af adresser beliggende på matrikler med en adresser. Der samme gør sig gældende for adresseprojektadressene. Her er 93,7 % af adresser beliggende på matrikler med flere adresser korrekt placeret, mens hele 98,7% af adresser med én adresse på matriklen er korrekt placeret. Begge testsandsynligheder er mindre end 0,001.

Antallet af lodder på matriklen har ligeledes meget stor betydning for, om adressen lander på det korrekte lod. Henholdsvis 95,9 og 98,6 % af adresserne beliggende på matrikler med ét lod er korrekt placeret, mens 68,8 % af stikprøveadresserne og 79,4 % af adresseprojektadresserne, som er beliggende på matrikler med flere lodder er korrekt placeret. Begge testsandsynligheder er mindre end 0,001.

Ser man bort fra adresseprojekt adresser uden klassifikation (3,7 %), er adresser beliggende i byen i højere grad korrekt placeret end adresser på landet. 97,8% af adresserne i byen er korrekt placeret mod 97,1% af landadresserne.

For stikprøveadresserne er der ikke signifikant forskel på land- og byadresser med hensyn til korrekt placering af adressen.

For Adresseprojektadresserne er der ligeledes undersøgt, om det har betydning for korrekt placering, om adressen er bebygget eller ubebygget. Konklusionen er klar: 95,1 % af de bebyggede adresser er korrekt beliggende, mens det kun er 82,1 % af de ubebyggede adresser der er korrekt beliggende.

### 6.3 anbefalinger til manuel årsagsforklaring

Den manuelle årsagsforklaring udføres af KMS. I det følgende gives anbefalinger til udvælgelsen af adresser til denne undersøgelse.

#### 6.3.1 Målsætning 1

*Stikprøveadresser.* Der er 26 stikprøveadresser, der ikke kunne genfindes i FLAT. Det anbefales, at et antal af disse adresser undersøges manuelt, for at kunne fastlægge årsagerne til den manglende eksistens i FLAT.

*Adresseprojektadresser.* 626 adresseprojekt adresser kunne ikke genfindes i FLAT, mens 1671 FLAT-adresser ikke havde nogen tilsvarende adresse i adresseprojektet.

Blandt de 626 adresser findes flere adresser med ens vejkode: 29% af de adresser der ikke kunne genfindes i kommune 217 er adresser med 5 eller flere manglende adresser indenfor samme vejkode. Dette drejer sig om 45% i kommune nr. 461, 93% i kommune nr. 653 og 0% i kommune nr. 821. En af årsagerne til, at de 626 adresser ikke kunne genfindes i FLAT kan således være, at det er nye adresser, som ikke indgår i grundlaget for FLAT.

Det anbefales derfor, at denne hypotese om forskel i udtrækstidspunkt afprøves i kommune 217, 461 og 653 for et antal adresser med vejkode, hvor mere end 5 adresser mangler. Blandt de øvrige adresser herunder alle i kommune 821 udvælges et antal for at finde andre årsagsforklaringer.

Af de 1671 FLAT adresser, der ikke kunne genfindes i adresseprojektadresserne er en meget stor del adresser med nklasse 'U'. Det vil sige adresser, der som følge af FLAT produktionsmetoden ikke er stedfæstet. I kommune 821 er halvdelen af de manglende adresser med NKLASSE 'U'.

Derudover er der en meget stor andel adresser, med samme vejkode. Mellem 45% og 73% af de manglende FLATadresser har vejkode, for hvilke mindst 5 adresser mangler i adresseprojektadresserne.

Det anbefales, at der udvælges et antal adresser med nklasse 'U' til manuel gennemgang. Endvidere anbefales at der udvælges adresser fra vejkode med mere end 5 manglende adresser. Endelig bør der udvælges et antal adresser blandt de øvrige adresser.

#### 6.3.2 Målsætning 3

For *stikprøveadresserne* er der 123 adresser, der ikke er placeret på korrekt lod-id. Det anbefales, at et antal blandt disse undersøges manuelt for årsager.

For *adresseprojektadresserne* er der 3627 adresser, som ikke er placeret på korrekt lod-id. Heraf ligger 3588 med en afstand på under 1000 m relativt tæt på den rigtige placering. Et antal af disse bør udtages for at undersøge om der er en systematisk årsag til at adresser placeres på forkert lod og om metoden hermed kan forbedres. Endvidere bør et antal blandt de 39 adresser med en afstand på mere end 1000 meter undersøges for at klarlægge årsager til at adresser kan placeres relativt langt fra den rigtige placering.

## 7. Konklusion

Der er lagt et stort arbejde i at udarbejde en produktionsmetode, der skal generere det bedst mulige resultat for stedfæstelse af adressedata ved samkøring af forskellige registre. De i denne rapport fundne resultater bærer da også tydeligt præg af denne grundighed, tabel 1. De fundne resultater er for nogle måls vedkommende væsentligt bedre, mens de for andre er knap så gode som de opstillede forventninger.

Det vil kunne forventes at 97,6% af alle adresser på landsplan vil findes i FLAT. De kommuner, for hvilke der i dag ikke eksisterer FLAT-data, er bykommuner. Da eksistensen er højere for byadresser end for landadresser, kan der forventes en endnu højere eksistens, såfremt de resterende kommuner havde indgået i vægtningen. Det kan overvejes at undersøge, om en optimering af den tidsmæssige forskydning på FLAT og grunddata kan bringe fuldstændigheden op på målsætningens 99%.

En umiddelbar svaghed ved FLAT er, at der er relativt mange adresser, som ikke umiddelbart kan genfindes i adresseprojektadresserne. Også her kan den manuelle årsagsbehandling måske medvirke til klarlægge, hvor der kan sættes ind for at nå målsætningen.

Det er desuden konstateret at testdata har indeholdt administrative adresser, som kan have negativ indflydelse på målopfyldelsen.

Selv om FLAT-metoden bygger på et princip om, at man hellere vil undvære en stedfæstelse end have en, der med stor sandsynlighed er forkert, er der alligevel 96,3% af adresserne, der er stedfæstede. En undersøgelse af, hvordan de ikke stedfæstede adresser fordeler sig på årsagsfaktorerne, kan medvirke til at klarlægge, på hvilke områder der skal sættes ind, for at nå kvalitetsmålet på 98%.

Et relativt stort antal adresser var placeret på forkert lod i forhold til kvalitetsmålene. Det gælder som ventet i særlig grad adresser med flere lodder eller adresser pr matrikel samt ubebyggede lodder. Men selv blandt de bebyggede lodder, lever FLAT ikke op til kvalitetsmålene. Det bør overvejes at undersøge hvor mange af de ”ukomplicerede” adresser (M1), der er placeret på korrekt lod-id, da dette vil være et udtryk for kvaliteten af produktionsgrundlaget.

Afstanden til den korrekt placerede adresse er meget lille i forhold til målsætningerne. Dette gælder særligt for de korrekt placerede adresser, men selv blandt de adresser, der ikke ligger på korrekt lod, er adressen placeret tæt på den korrekte stedfæstelse.

Den foreliggende undersøgelse har givet en grundlæggende beskrivelse af FLAT adressernes kvalitet. Undersøgelsen skelner imidlertid ikke mellem kvaliteten af adresser med forskellige karakteristika. En dybere analyse af kvaliteten og produktionsmetoden kan bl.a. omfatte:

- Hvor stor en andel af uoverenstemmelserne skyldes kommunernes administrative adresser?
- Hvilken kvalitetsforskel er der på beboede og ubeboede adresser?
- Hvor høj en andel af adresser placeret på forkert matrikelod er placeret på en anden af ejendommens lodder i umiddelbar nærhed af den korrekte placering? Hvad er kvaliteten af disse adresser i forhold til adresser placeret på korrekt lod?
- Hvilken sammenhæng er der mellem produktionsmetodens klassifikation og nøjagtighed i stedfæstelse?
- Hvilken nøjagtighed i afstand kan man forvente for adresser med stigende antal adresser eller lodder pr matrikel?

Endelig kan det overvejes, om metoden er for restriktiv i forbindelse med fjernelse af stedfæstede adresser (supplerende kontrolrutiner). Særligt kontrol 2, hvor en stedfæstelse fjernes, hvis der på grundlag af TOP10DK bygninger findes en kandidat tættere på naboen end den oprindelige kandidat, er

interessant i denne sammenhæng. I denne sammenhæng bør undersøges hvor mange adresser, der får fjernet stedfæstelsen og hvad kvaliteten af den stedfæstelse der blev fjernet var.

## 8. Bilag 1: Vægtning af stikprøve-adresser

Fordelingen af stikprøve-adresser på en række nøglefaktorer (en eller flere adresser på matriklen, en eller flere lodder på matriklen og om adressen ligger i land eller by) er undersøgt i forhold til hele landet.

Fordelingen af landets adresser på disse typer<sup>1</sup> kan ses af nedenstående tabel

**Adresstype Hele landet (FLAT-adresser)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 adr by et lod	982067	49,4	53,7	53,7
	1 adr land et lod	388593	19,5	21,3	75,0
	1 adr by flere lod	8691	,4	,5	75,4
	1 adr land flere lod	32343	1,6	1,8	77,2
	flere adr by et lod	282454	14,2	15,4	92,7
	flere adr land et lod	69966	3,5	3,8	96,5
	flere adr by flere lod	42698	2,1	2,3	98,8
	flere adr land flere lod	21551	1,1	1,2	100,0
	Total	1828363	91,9	100,0	
Missing	System	160734	8,1		
Total		1989097	100,0		

Der er beregnet vægte på baggrund af disse data, som fordeler sig som følger:

Adresse type	Faktor
1 adr by et lod	592,68
1 adr land et lod	533,05
1 adr by flere lod	724,25
1 adr land flere lod	385,04
Flere adr by et lod	155,95
Flere adr land et lod	885,65
Flere adr by flere lod	1423,27
Flere adr land flere lod	1077,55
Mangler værdier	1071,56

For adresser, der mangler oplysninger om en eller flere af typevariablene er der ikke skelnet. Der er således beregnet en samlet vægt for disse uafhængig af, om der mangler oplysninger om få eller alle typevariable.

<sup>1</sup> Se afsnit 5.3

## 9. Bilag 2: Vægtning af Adresseprojektadresser.

Der er beregnet vægte på baggrund af disse data, som fordeler sig som følger:

Adresse type	Faktor
1 adr by et lod bebygget	15,53
1 adr land et lod bebygget	36,37
1 adr by flere lod bebygget	38,96
1 adr land flere lod bebygget	45,42
Flere adr by et lod bebygget	11,57
Flere adr land et lod bebygget	19,72
Flere adr by flere lod	10,30
Flere adr land flere lod	41,93
1 adr by et lod ubebygget	26,62
1 adr land et lod ubebygget	34,19
1 adr by flere lod ubebygget	14,87
1 adr land flere lod ubebygget	25,95
Flere adr by et lod ubebygget	15,12
Flere adr land et lod ubebygget	24,60
Mangler værdier	81,38

For adresser, der mangler oplysninger om en eller flere af typevariablene er der ikke skelnet. Dog er adresser der stammer fra matrikler med flere adresser og flere lodder ikke opdelt efter by og land. Der er således beregnet en samlet vægt for disse uafhængig af, om der mangler oplysninger om få eller alle typevariable.

## 10. Bilag 3 Afstand til korrekt placering

Tabel 1: Afstand af adresser. Vægtede data. Tal i parentes er realiserede data.

ADRESSER I ALT (KUN ADRESSER, DER KUNNE GENFINDES I FLAT)	BY-ADRESSER			LANDADRESSER			ADRESSER MED BÅDE BY OG LAND	IKKE KLASSIFICERDE ADRESSER
		Heraf sted- fæstede	Indenfor en afstand af 100m		Heraf sted- fæstede	Indenfor en afstand af 250m		
1941483 (107832)	1315519 (90942)	1314806 (90896)	1238645 (84215)	511783 (15487)	510697 (15457)	502068 (15155)	22299 (274)	91883 (1129)