

Oprensning af vægtdata - statistikrapport	Ansvarlig	AMK/TRIA
	Oprettet	10-12-2021
Projekt: 5517 Datadrevetmanagement – AP1	Side	1 af 10

STØTTET AF  
**mælkeafgiftsfonden**

## Oprensning af vægtdata fra automatiske malkesystemer

### Sammenfatning

I denne rapport beskrives 2 forskellige metoder til at oprense/korrigerer vægte fra forskellige robotter, så de let kan anvendes i produktionsstyringsværktøjer. De to metoder bygger på to forskellige antagelser:

- Metode A: Oprensning med en mixed model – som bygger på en antagelse om, at køer med samme race, laktationsstadiet og laktationsnumre vejer det samme uanset, hvilket hold og hvilken robot, de går ved.
- Metode B: Oprensning, som bygger på en antagelse om, at køer der indenfor kort tid er malket ved forskellige robotter, skal veje det samme.

Ulempen ved metode B er, at der i nogle besætninger ikke flyttes køer mellem alle robotter særligt ofte. Ulempen ved metode A er, at antagelsen om, at køer på samme laktationsstadiet og laktationsnummer vejer det samme, ikke altid holder, f.eks. hvis der fodres forskelligt på de forskellige hold eller hvis køerne er opdelt efter størrelse. Til gengæld er metode A mere robust end metode B.

En opgørelse viser stor korrelation mellem vægte oprenset ved de to metoder. I de fleste besætninger var korrektionen mellem robotter kun lille og i langt de fleste situationer har oprensningen fungeret, og det er muligt at anvende de korrigerede vægte i dataanalyse af køers vægt og i produktionsstyringen. I enkelte tilfælde kunne der dog i perioder være over 100 kg forskel mellem robotter, og der bør derfor gives en tilbagemelding til kvægbrugeren vedr. robotforskelle, hvis man vil anvende vægtene til produktionsstyring.

### Indledning

Nogle robotter vejer automatisk køerne ved hver malkning. I besætninger er der ofte flere robotter, der kan veje forskelligt, selv om robotterne jævnlige kalibreres. Det kan være et problem, hvis vi vil bruge vægtene fra robotterne til produktionsstyring, derfor beskriver denne rapport 2 måder at automatisk oprense vægtdata på. Oprensningen kompliceres af, at der findes mange forskellige systemer at inddele køerne i hold på, og dermed også hvilke robotter, de forskellige køer går til. Derudover kalibreres robotterne jævnlige, uden at vi har oplysninger om hvornår, så vi kan ikke regne med, at forskellen er konstant.

### Detaljeret beskrivelse af metoderne til oprensning

Lige meget om man bruger metode A eller B startes der med en oprensning af åbenlyse fejl. Dvs.:

- Ko-id manglende
- Besnr manglende
- Starttid og sluttid for besøg manglende
- CKR. Dyr. Nr. manglende
- Dato for besøget start er efter trækningstidspunktet
- Dato for besøget start er før d. 1. januar 2002.
- Dobbelt observationer sorteres fra. Dvs. besøg med samme besøgstider i samme robot med samme mælkeydelse.
- Jerseykøer med vægte under 100 kg eller over 900 kg
- Ikke Jerseykøer med vægte under 300 kg eller over 1100 kg

### **Oprensning metode A – ens vægt uafhængig af robot**

Efter den indledende grove oprensning, som beskrevet ovenfor, køres en normal lineær model på de groft oprensede data for at korrigere for forskelle i de køer, der bliver malket ved forskellige robotter. Modellen ser ud som følgende:

$$BW = \text{aar\_uge} + \text{aar\_uge} * \text{robot\_id} + \text{race\_J} + \text{klvnr} (1., 2., \text{øvrigte}) + \text{race\_J} * \text{klvnr} * \text{DIM} + \text{klvnr} * \text{DIM}^2 + \text{klvnr} * \text{DIM\_W}$$

Hvor BW står for vægten af koen, aar\_uge er en effekt af uge inden for år, robot\_id er effekten af robotten, race\_J er en effekt af om koen er en Jersey ko eller ikke, klvnr er en effekt af om koen er i 1., 2. eller 3. og øvrige laktationer, DIM er dage efter kælvning og DIM\_W står for effekten af  $\exp(-0.10 * \text{DIM})$  – eksponentialfunktion på dage efter kælvning.

Derefter udregnes robotkorrektionen ud fra de estimerede effekter som:

Robotkorrektion=effekt af robot i uge\_xx/år\_xx – den gennemsnitlige effekt af alle robotterne i besætningen i uge xx/år xx

Udregningen sker inden for uge/år, fordi det er prioriteret højt ikke at komme til at korrigere evt. tidsmæssige ændringer i vægt i hele besætningen ud, men til nogle anvendelser ville det måske være mest relevant at lave en samlet korrektion og ikke en korrektion pr. uge. Dette kan gøres ved i stedet for at tage gennemsnit pr. uge/år at tage et samlet gennemsnit af roboteffekterne og trække det fra.

Hvis spredningen i de robotkorrigerede vægte for en robot i en uge er over 200 eller under 1 bruges resultaterne for robotten ikke.

For hver ko indenfor laktation bortsorteres outlier. Det sker hvis en kos vægt afviger mere end 3,5 gange spredningen på koens vægte fra koens median vægt i den laktation. Der bruges kun vægte fra køer med mere end 5 vægte.

Til sidst findes en udglattet vægt ved hjælp af proceduren proc loess i SAS med standardopsætning.

### **Oprensning metode B – ens vægt ved malkning ved 2 forskellige robotter inden for kort tid**

Her bruges data kun, hvis koen har mere end 10 malkninger pr. laktation pr. robot.

Først udelukkes grove outlier som beskrevet tidligere i dette dokument. For hver ko pr. laktation pr. robot. udglattes de oprensede vægtene derefter ved hjælp af proceduren proc loess i SAS med standardopsætning. Der bruges kun observationer, hvor den reelle vægt afviger mindre end 150 kg fra den udglattede vægt.

Derefter dannes en ny udglattet vægt på samme måde ud fra vægtene uden de groveste outlier. For hver ko pr. laktation pr. robot. udglattes de oprensede vægte ved hjælp af proceduren proc loess i SAS med standardopsætning. Igen bruges kun køer med mere end 10 observationer pr. ko pr. robot.

For alle køer, der inden for 48 timer er malket ved 2 forskellige robotter, findes derefter forskellen i de udglattede vægte fra de to robotter.

Til udregning af roboteffekterne pr. uge bruges alle resultaterne fra køer, der er malket ved to forskellige robotter fra den sidste uge. Hvis der er mindre end 25 "flytninger" mellem to robotter, bruger man dog også ældre data, indtil man har 25 "flytninger", hvor køer er malket ved de to malkerobotter. Dvs. at der godt kan indgå væsentlig ældre data end fra den sidste uge i robotkorrektionen.

Derefter udregnes roboteffekterne pr. uge i en model, hvor forskellen i vægt mellem robotterne er responsvariabel, og de forklarende variabler er hvilke robotter, køen er malket ved, og i hvilken rækkefølge den er malket.

Igen findes robotkorrektionen som

Robotkorrektion=effekt af robot i uge\_xx/år\_xx – den gennemsnitlige effekt af alle robotterne i besætningen i uge xx/år xx.

Hvis det ikke er muligt at beregne robotkorrektion, pga. at der er for få køer, der er malket ved forskellige robotter, bruges den ukorrigerede vægt.

Ud fra de korrigerede vægte findes derefter outlier. Det sker ved, at der for hver ko udregnes et løbende gennemsnit ud fra følgende formel:

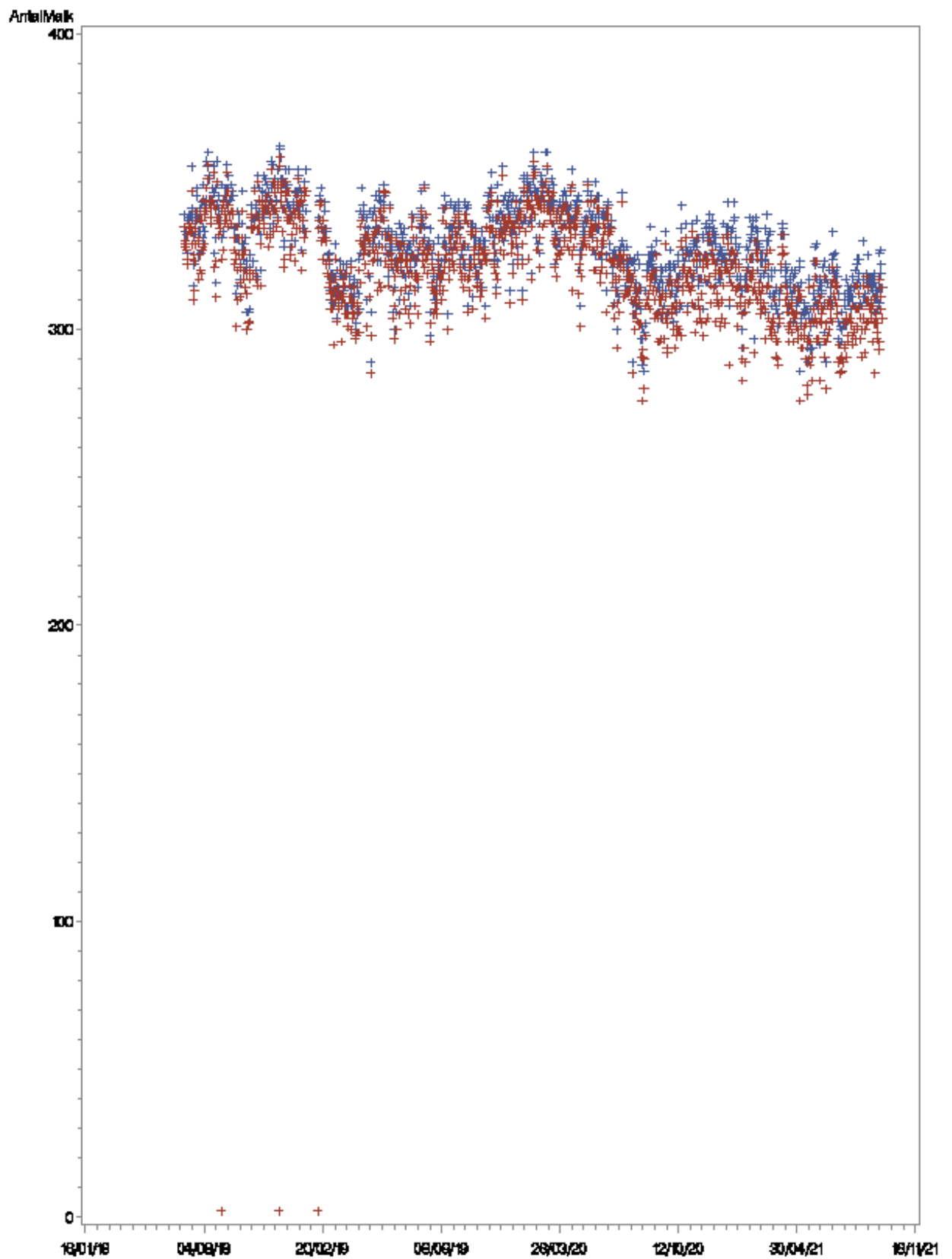
Løbende gennemsnit= $0,2 \cdot \text{Sidste vægt} + 0,8 \cdot \text{sidste løbende gennemsnit}$

Som startværdi for det løbende gennemsnit bruges medianen af de første 6 målinger. Der bruges en ny startværdi ved opstart af en ny laktation, og/eller hvis der har været mere end 30 dages pause i målingerne.

En måling markeres som outlier, hvis den afviger meget fra det løbende gennemsnit. Med meget menes over 60 kg eller 2,5 gange spredningen på afvigelsen fra det løbende gennemsnit, hvis spredningen er over 24 kg.

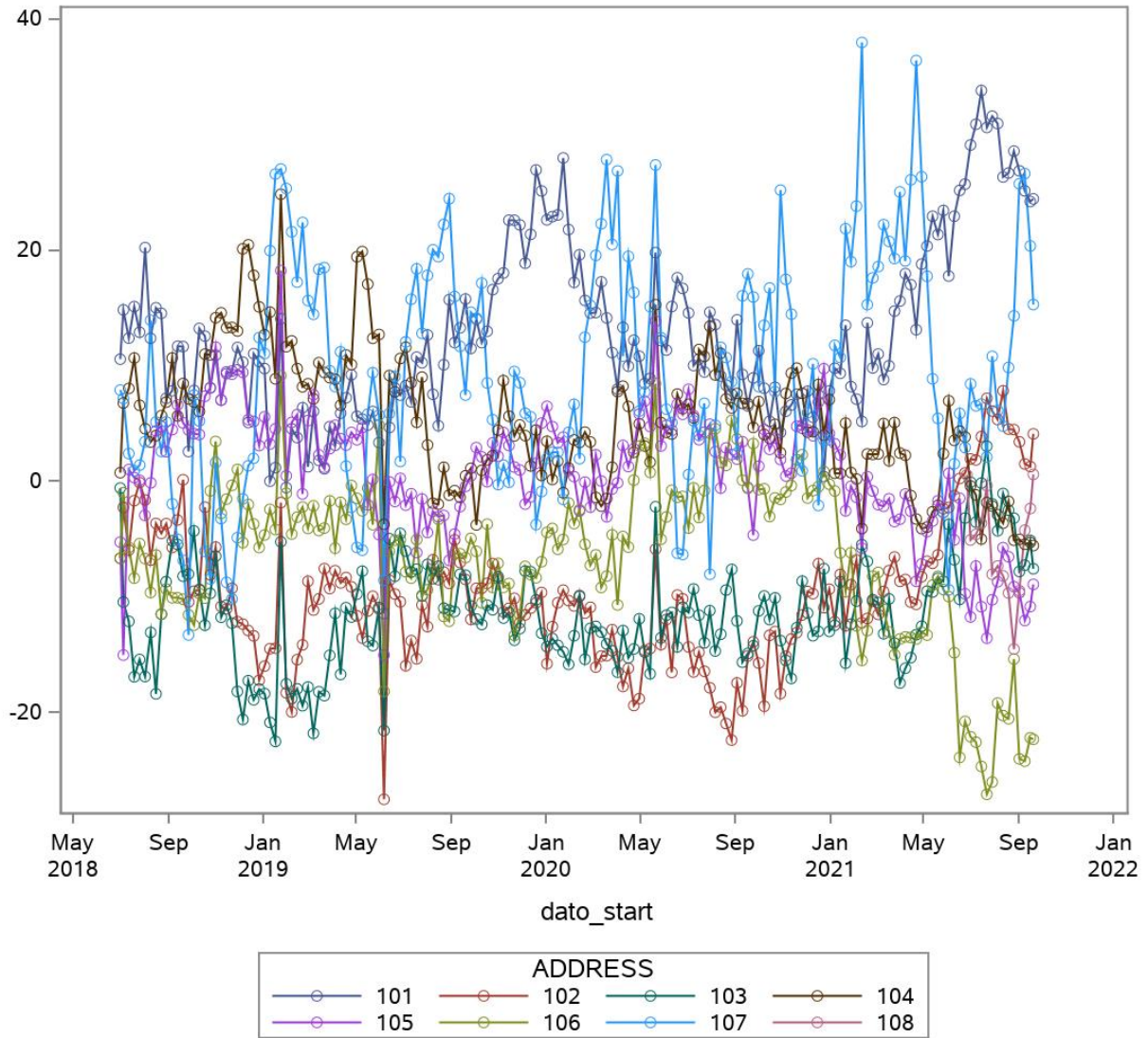
### **Effekten af oprensning**

For hver besætning, hvor data er oprenset, dannes en række figurer/tal, som man kan bruges til at se, hvordan oprensningen har fungeret. I figur 1-5 er vist eksempler på, sådanne figurer.

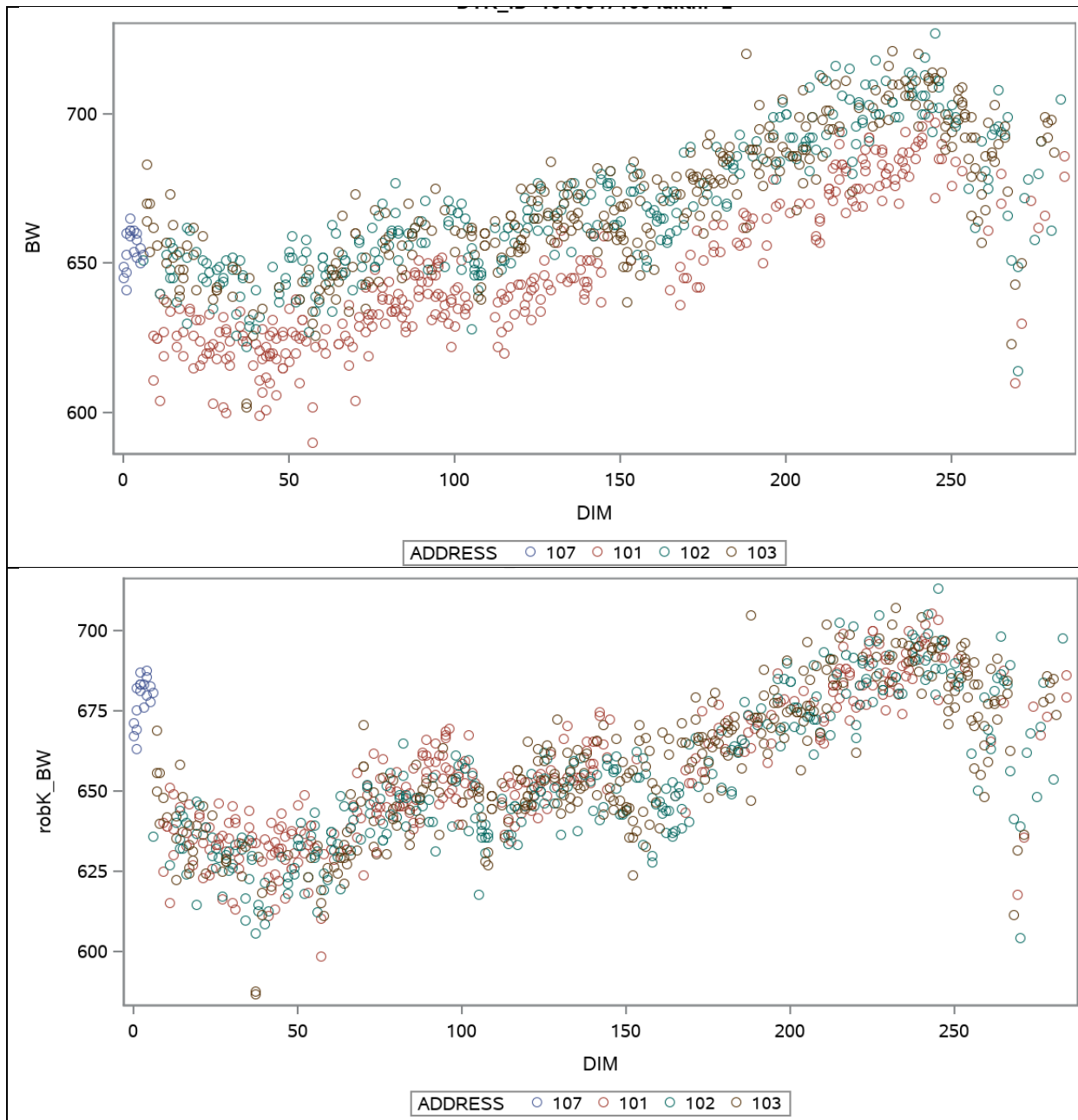


Figur 1. Oversigt over udviklingen i antal malkninger og antal vejninger pr. dag i en tilfældig besætning. Bemærk hullet i antal overførte malkedata fra robotterne i perioden fra d. 23/1 2019 til d. 12/2 2019.

## Robotkorrektion

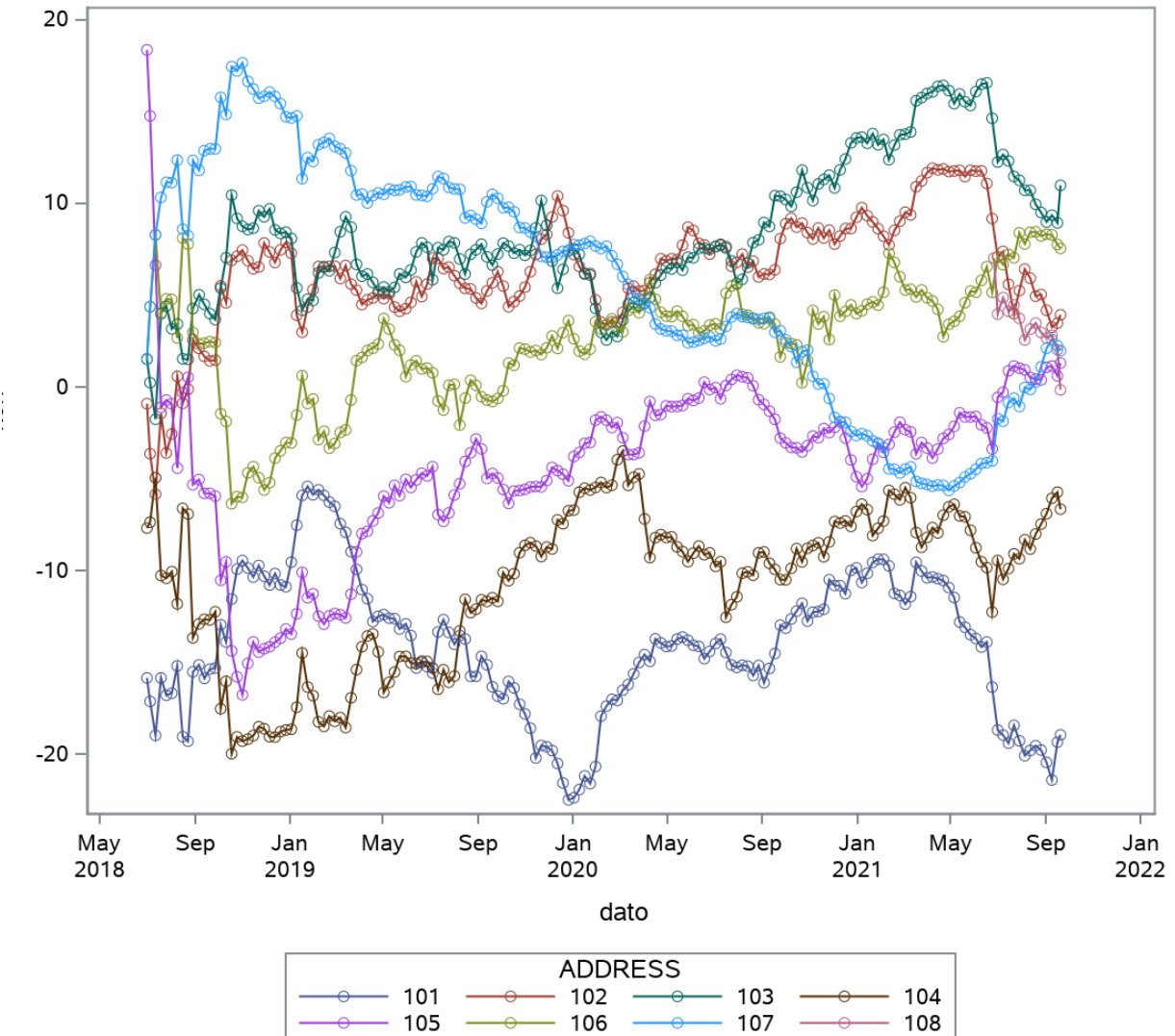


**Figur 2. Eksempel på udviklingen i robotkorrektionen ved metode A i en tilfældig besætning. NB. Samme besætning som figur 4.**

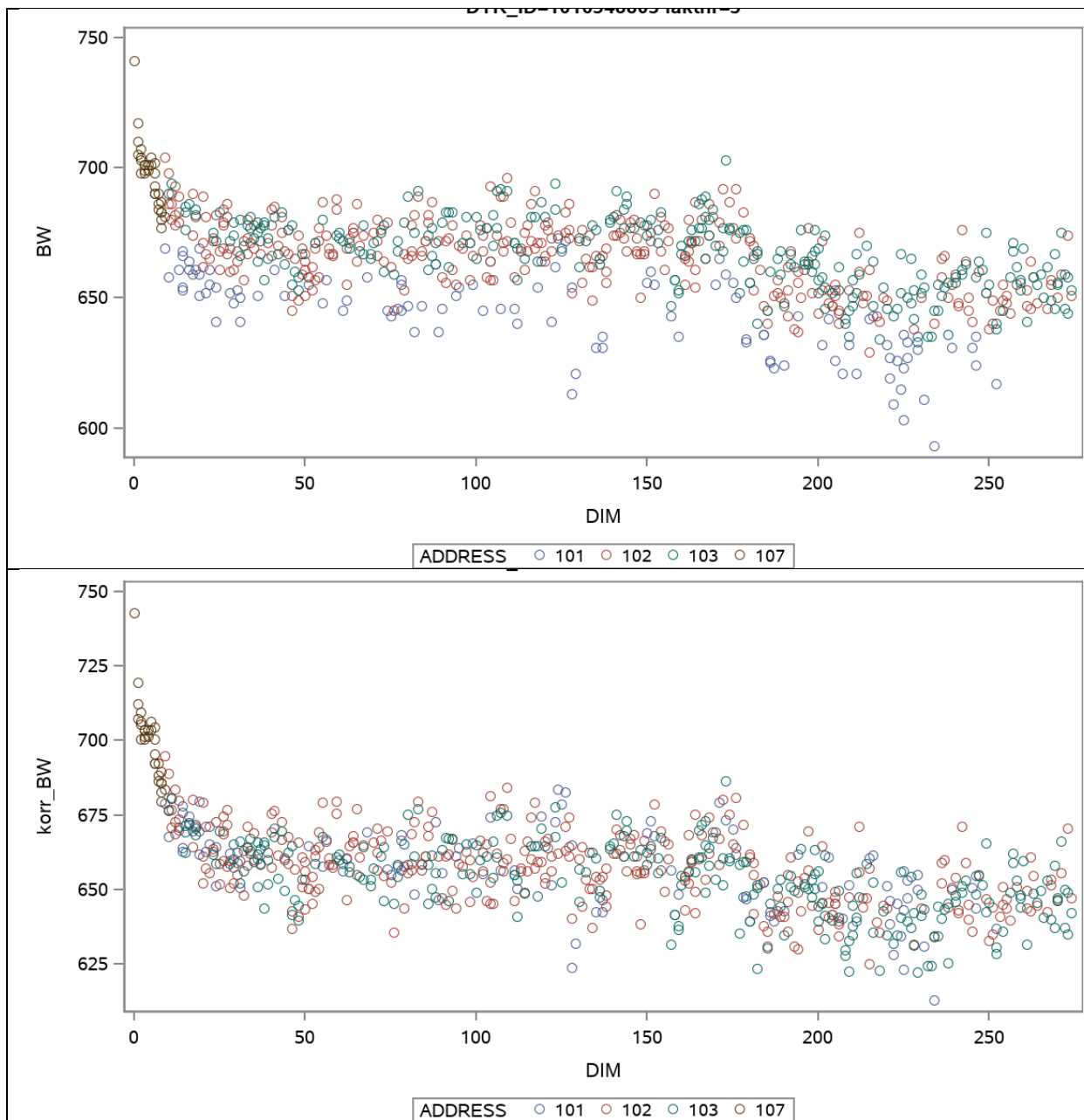


**Figur 3. Effekten af robotkorrektionen ved metode A for en tilfældig ko. Øverste billede viser vægtene før korrektion, og nederste billede viser vægtene efter korrektion.**

### Robotkorrektion, sammenvejet



**Figur 4. Oversigt over udviklingen i robotkorrektionen ved metode B i en tilfældig besætning. NB. Samme besætning som figur 2.**



**Figur 5. Effekten af robotkorrektionen ved metode B for en tilfældig ko. Øverste billede viser vægtene før korrektion, og nederste billede viser vægtene efter korrektion.**

Ud over figurerne dannes der også forskellig tal til vurderingen af oprensningen. I tabel 1 er vist korrelationen mellem de tre robotkorrigerede vægte:

- Korrigerede vægte metode A
- Udglattede korrigerede vægte ved metode A
- Korrigerede vægte metode B

I tabel 1 ses, som eksempel, resultater fra de 15 første besætninger. For langt de fleste besætninger er der meget høj korrelation mellem de tre korrigerede vægte. Ud af 124 besætninger er der kun 9 besætninger, hvor korrelation mellem de korrigerede vægte ved metode A og B er mindre end 0.95, der er 11 besætninger, hvor korrelation mellem de udglattede vægte metode A og de korrigerede vægte metode B er mindre end 0.95 og der er 5 besætninger, hvor korrelation mellem de korrigerede vægte A og de udglattede vægte metode A er mindre end 0.95. Ud af dem er der 3 besætninger, hvor korrelationerne er nul/manglende, fordi den ene af de korrigerede vægte mangler. Derudover er korrelationerne meget lave



for en besætning. Men generelt ser robotkorrektionen ud til at virke fornuftigt, og for langt de fleste besætninger er resultatet uafhængig af metode.

Tabel 1. Korrelation mellem vægte oprenset ved forskellige metoder for de 15 første besætninger, som kun er et udsnit af de 124 besætninger som indgår i undersøgelsen.

Be-sæt-ning	Korrelation metode A og B korrigerede vægte	Korrelation metode A udglattet og metode B korrigerede vægte	Korrelation metode A korrigerede vægte og udglattede korrigerede vægte
1	0.98	0.96	0.98
2	0.99	0.98	0.99
3	0.99	0.98	0.99
4	1.00	0.99	0.99
5	1.00	0.99	0.99
6	0.99	0.98	0.99
7	1.00	0.98	0.98
8	1.00	0.98	0.98
9	1.00	0.99	0.99
10	1.00	0.99	0.99
11	1.00	0.99	0.98
12	0.99	0.99	0.99
13	0.97	0.96	0.99
14	0.99	0.98	0.99
15	0.68	0.63	0.88

I tabel 2 er vist gennemsnit af spredning på de tre korrigerede vægte pr. ko pr. laktation. for de fleste besætninger er spredningen på de udglattede vægte som forventet mindre end på de rent korrigerede vægte, mens der ikke er stor forskel på spredningen mellem de korrigerede vægte udregnet ved de to metoder. For nogle enkelte besætninger er den gennemsnitlige spredning stor (over 50 kg) - det drejer sig om 4 besætninger ud af de 124 besætninger. For resten af besætningerne ser det fornuftigt ud.

Tabel 2. Gennemsnit af spredning på korrigerede vægte efter forskellige metoder for de 15 første besætninger. Spredningen er udregnet pr. ko pr. laktation.

Besæt-ning	Antal køer	Spredning på korrigerede vægte metode A	Spredningen på korrigerede og udglattede vægte metode A	Spredningen på korrigerede vægte metode B
1	916	28	23	31
2	1980	30	28	30
3	819	31	27	32
4	290	27	24	27
5	269	13	8	14
6	609	32	30	32
7	263	13	7	13
8	190	30	26	30
9	716	24	21	24
10	915	29	25	29
11	502	31	27	31
12	675	30	27	30
13	1213	34	32	37
14	658	36	33	39
15	1837	57	34	65

I tabel 3 er vist fordelingen af forskellen mellem de korrigerede vægte og de rå vægte for udvalgte besætninger. Der undersøges forskellen mellem den målte vægt og hver af de tre korrigerede vægte

- R1 forskellen mellem den robotkorrigerede vægt ved metode A og den rå vægt
- R2 forskellen mellem den udglattede og robotkorrigerede vægt ved metode A og den rå vægt
- R3 forskellen mellem den robotkorrigerede vægt ved metode B og den rå vægt

Tabel 3. Fordelingen af forskellen mellem de korrigerede vægte og de rå vægte for de 15 første besætninger. P5 står for 5 % fraktilen, P95 står for 95 % fraktilen og min står for minimum og max står for maksimum.

Be-sæt-ning	R1 P5	R2 P5	R3 P5	R1 P95	R2 P95	R3 P95	R1 Min	R2 Min	R3 Min	R1 Max	R2 Max	R3 Max
1	-14	-28	-3	12	28	5	-78	-145	-8	127	256	6
2	-16	-23	-12	19	24	17	-28	-365	-18	38	358	22
3	-11	-22	-8	15	23	6	-36	-254	-18	34	326	11
4	0	-15	-2	0	16	6	0	-218	-5	0	344	8
5	-7	-16	-4	5	16	5	-9	-83	-5	6	79	5
6	-14	-22	-4	15	21	4	-19	-308	-5	19	302	5
7	-4	-16	-2	6	17	3	-8	-53	-3	6	52	3
8	-58	-65	-58	58	64	58	-60	-162	-60	60	195	60
9	-5	-15	-1	5	16	1	-30	-228	-4	8	99	4
10	-7	-18	-4	7	20	5	-12	-296	-10	16	162	15
11	-9	-20	-8	9	20	8	-14	-330	-14	14	581	14
12	-15	-22	-5	17	23	5	-21	-340	-11	26	204	18
13	-26	-33	-22	26	31	14	-38	-321	-28	43	254	38
14	-95	-86	-83	51	50	38	-106	-313	-98	60	352	80
15	-352	-351	-346	192	201	224	-502	-624	-446	328	556	394

I langt de fleste besætninger er der mindre end 20 kg i forskel både for P5 og for P95. I 15 af besætningerne er forskellen dog over 50. Et ekstremt tilfælde er besætning nr. 15, hvor der korrigeres med ca. 350 kg. Generelt vil det i besætninger med høj korrektion være en god ide at melde tilbage til kvægbrugeren, at der er behov for kalibrering, hvis vægtene ønskes anvendt til produktionsstyring.

### Samlet vurdering af oprensningen

Der er undersøgt 124 besætninger. For langt de fleste af disse besætninger har oprensning og robotkorrektion virket efter hensigten, og det dermed er muligt at estimere en brugbar kropsvægt. For enkelte besætninger er der udfordringer. Det kan eksempelvis skyldes, at besætningen har to robotter, som vejer meget forskelligt. Der bør derfor laves en kalibrering af vejecellerne, hvis landmanden ønsker brugbare vægte. En anden udfordring kan være, at spredningen er meget stor mellem kørerne i besætningen, og det kan igen skyldes brug af forskellige vægte, som ikke er kalibrerede korrekt. Der er også nogle besætninger, hvor der kan være en periode, hvor det ikke fungerer, hvorefter det igen fungerer.

Generelt er det svært at lave en oprensning og robotkorrektion, som passer perfekt på alle køer og i alle situationer, så det vil altid være nødvendigt at være kritisk overfor resultaterne og holde øje med udviklingen indenfor hver besætning. Men i langt de fleste situationer har oprensningen fungeret, og det er muligt at anvende de korrigerede vægte i dataanalyse af køers vægt og i produktionsstyringen.

### Forhold, der kan gøres bedre

Hvis man senere skal lave en forbedret oprensning, bør man være opmærksom på følgende forhold:

- Der er ingen udelukkelse af dage med unormal mælkeydelse på robotniveau. F.eks. kan perioder med ekstrem frost betyde, at nogle af vægtene vejer forskelligt.
- Der er ikke tjekket for om nogle robotter har ekstremt svingende resultater.
- Der er ikke tjekket på meget lange malketider, som f.eks. kan opstå ved overgangen til sommertid
- Der er ikke lavet nogen automatisk algoritme til at fastlægge hvilket hold, de enkelte køer går på, de enkelte dage. Det ville ud fra malkedataene være ret let at lave en sådan algoritme.