Factors Affecting the Accuracy of the Recording Day
Session 3a-b
Estimation of carryover in automatic milking systems based on farm tests by ICAR test centres

Christian Ammon, MelkFEE, Germany. Data scientist.

Milk sampling in an AMS during an ICAR farm test

- AMS and sampler combination

- Taking the device sample

- Taking the reference sample

- What can go wrong
  - Homogenisation
  - Carryover

Modelling carryover in an AMS - prerequisites

- Farms: usually two with at least one AMS box each
- Sampling devices depending on the requested combinations, usually two per combination
- 50 milkings per farm and day, ideally tracking cow IDs
- Device sample (usually one per milking)
  - All samples are the same size/volume
  - Sample volume does not depend on milk yield
- Reference samples (usually two per milking – calculate mean)
- Get milk sample data from the milk laboratory
  - Relevant for ICAR farm test: fat content
  - But we might be interested in additional components …

ICAR
THE GLOBAL STANDARD
FOR LIVESTOCK DATA

The (basic) statistical look at carryover

$$y_i = \mu + \beta_1 f_t + \beta_2 f_{t-1} + \varepsilon_i$$

- y_i: observed fat content in the sampler at ith milking
- μ: intercept
- β_1:   estimated regression coefficient for current reference fat content
- f_t: current reference fat content at milking t
- β_2:   estimated regression coefficient for previous reference fat content
- f_{t-1}:   previous reference fat content at milking t-1
- ε_i: random residual for ith milking, homoscedasticity, N(0,σ)
- Add random effects for farm, sampler, cow, repeated measures …

Example: carryover estimations based on ICAR farm test setups for different AMS

| Device	Farms	Boxes and samplers	Milkings	Cow effect included	Intercept	Reference		Previous reference		SP(POW)	Carryover (calculated)
						Estimated coefficient	Pr > F	Estimated coefficient	Pr > F		
A	1	2; 2	48; 45	YES (sim.)	n.s.	0.5965	<.0001	0.2674	0.0006	0.2185	31%
A	1	2; 2	48; 45	NO	n.s.	0.5996	<.0001	0.2754	<.0001	0.1201	31%
B	1	1; 2	50; 50	YES (sim.)	n.s.	0.9664	<.0001	0.02077	n.s.	0.1284	2%
B	1	1; 2	50; 50	NO	n.s.	0.9633	<.0001	0.02735	n.s.	0.1249	3%
C (1)	2	2; 1	67; 50	YES	n.s.	1.0153	<.0001	0.004603	n.s.	0.6302; 0.1442	0%
C (1)	2	2; 1	67; 50	NO	n.s.	1.0213	<.0001	-0.0007	n.s.	0.2392; 0.1451	0%
C (2)	2	2; 2	50; 51; 50; 50	YES	n.s.	0.9992	<.0001	0.003264	n.s.	0.1095; 0.1014	0%
C (2)	2	2; 2	50; 51; 50; 50	NO	n.s.	0.9992	<.0001	0.003265	n.s.	0.1095; 0.1014	0%

Can we trust the results?

- Are up to 200 milkings enough to reliably estimate carryover?
- Actually: 100 milkings per AMS box, or 50 milkings per AMS/sampler combination
- Are the results repeatable?
- Which range of carryover can I detect with that test setup?
- How accurately can I determine the carryover?

THE GLOBAL STANDARD
FOR LIVESTOCK DATA

Setting up a statistical simulation

- Create simulated milkings:
  - Generate list of cows with their milkings (also include farms or different samplers if necessary)
  - Create reference fat contents based on statistical distribution (random number from mean and standard deviation for fat in a herd)
  - Calculate device fat content by applying a carry-over effect: e.g. 0.85 times reference + 0.15 times previous reference results in new mean to randomize sample fat content from for a carry-over of 15 %
- Do this for your intended range of carry-over (e.g. 2% to 25 % in steps of 1 %)
- Run the statistical model with this simulated data set for a couple times, depending on capacity (I tried to get at least 200 runs in per variant)

Results for farm test simulation with fat content (Power)
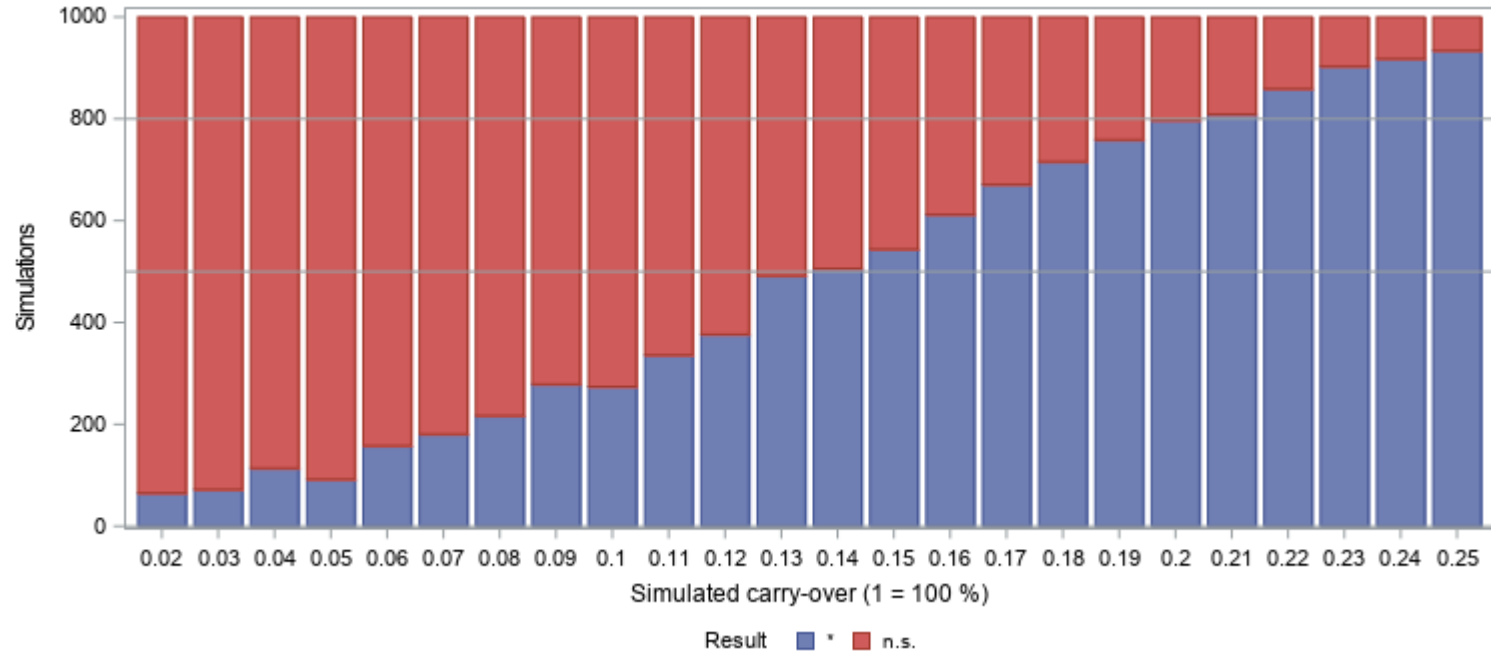
Results for farm test simulation with fat content (calculated carry-over)
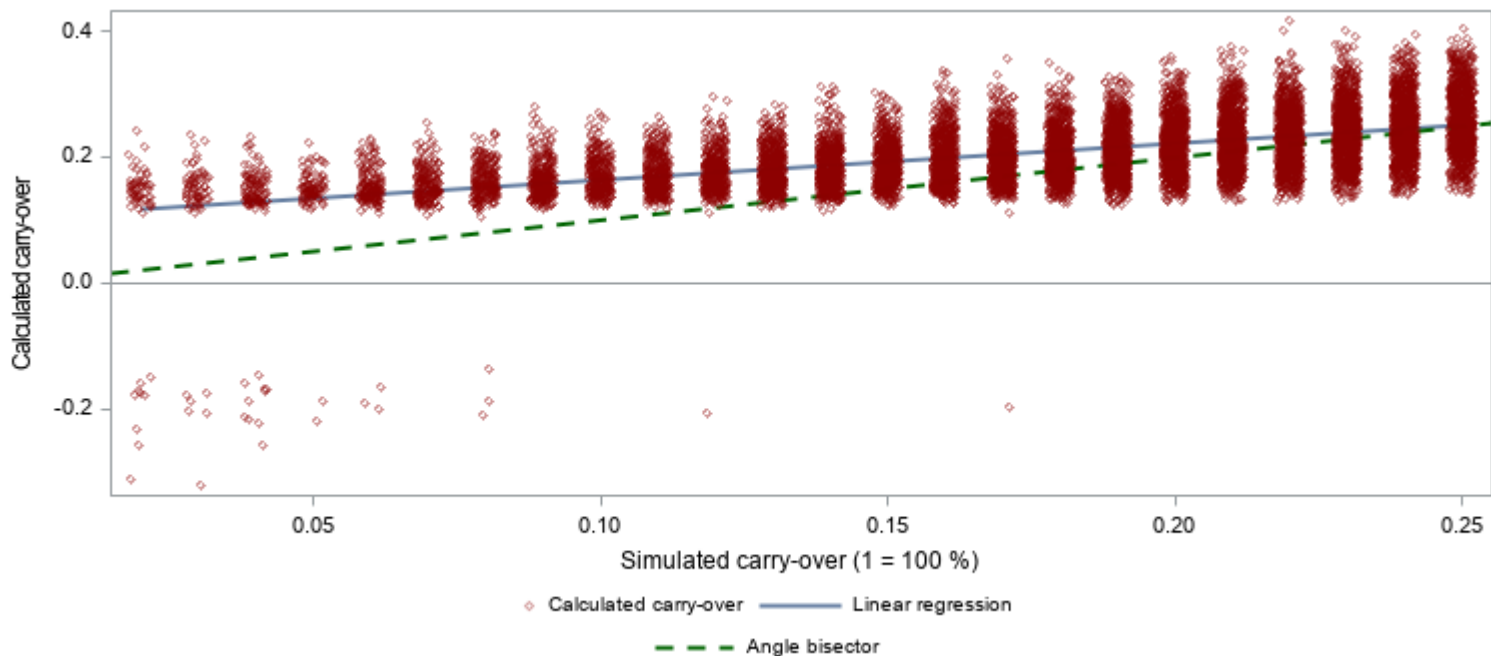
First conclusion

- Results are not looking great
- Reliable estimates for a significant influence cannot be expected for carry-over below 20 % in this setup
- Carry-over is overestimated for low actual carry-over, and there seems to be a pattern

- We need to try something else!

Using additional data

- Add data from other milk components (e.g. protein, lactose …)
- When AMS milk is well homogenized, carryover should be the same for all components
- Adjust statistical model to include more components (statistically not quite correct as these are no true repetitions, but we should get a good idea if the procedure improves)
- Updated formula introduces component c with *j* levels:

$$y_{ij} = \mu + c_j + \beta_1 \begin{pmatrix} f \\ p \\ l \end{pmatrix}_t + \beta_2 \begin{pmatrix} f \\ p \\ l \end{pmatrix}_{t-1} + \varepsilon_{ij}$$
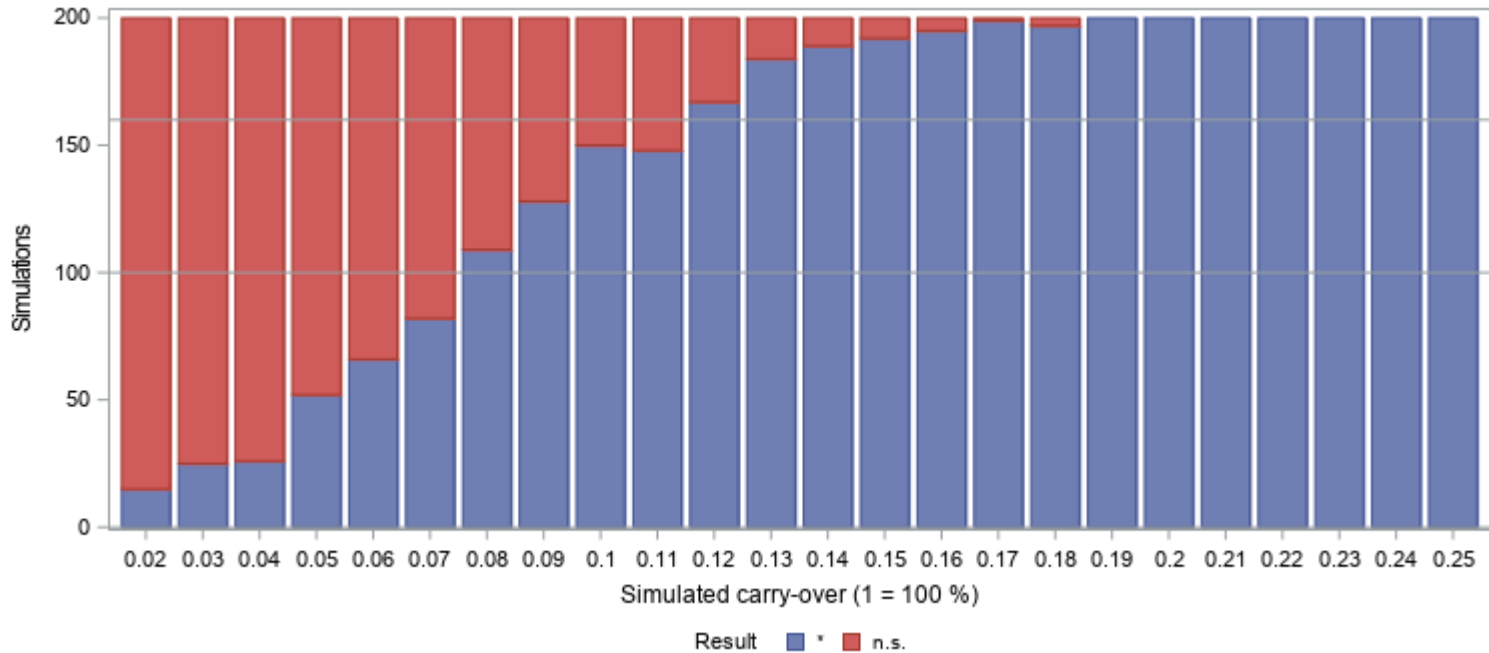
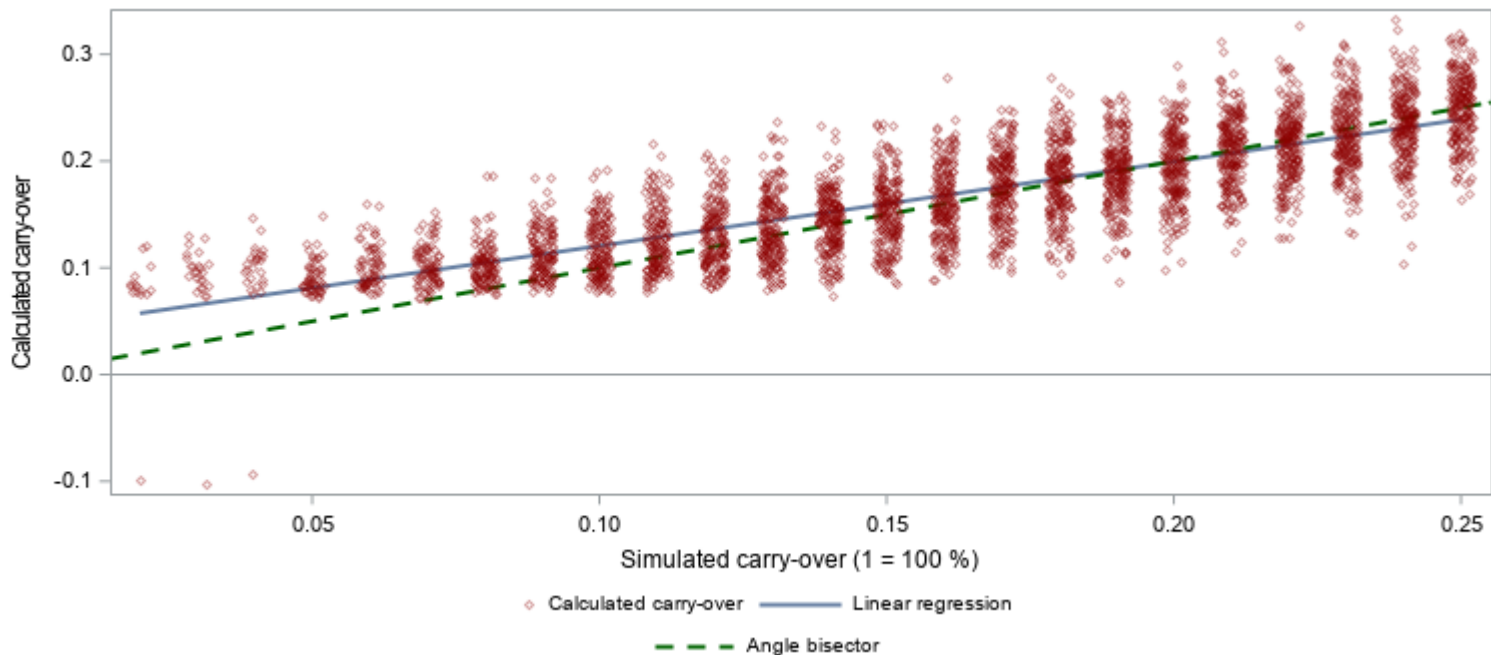# Results for farm test simulation with fat and protein contents (Power)
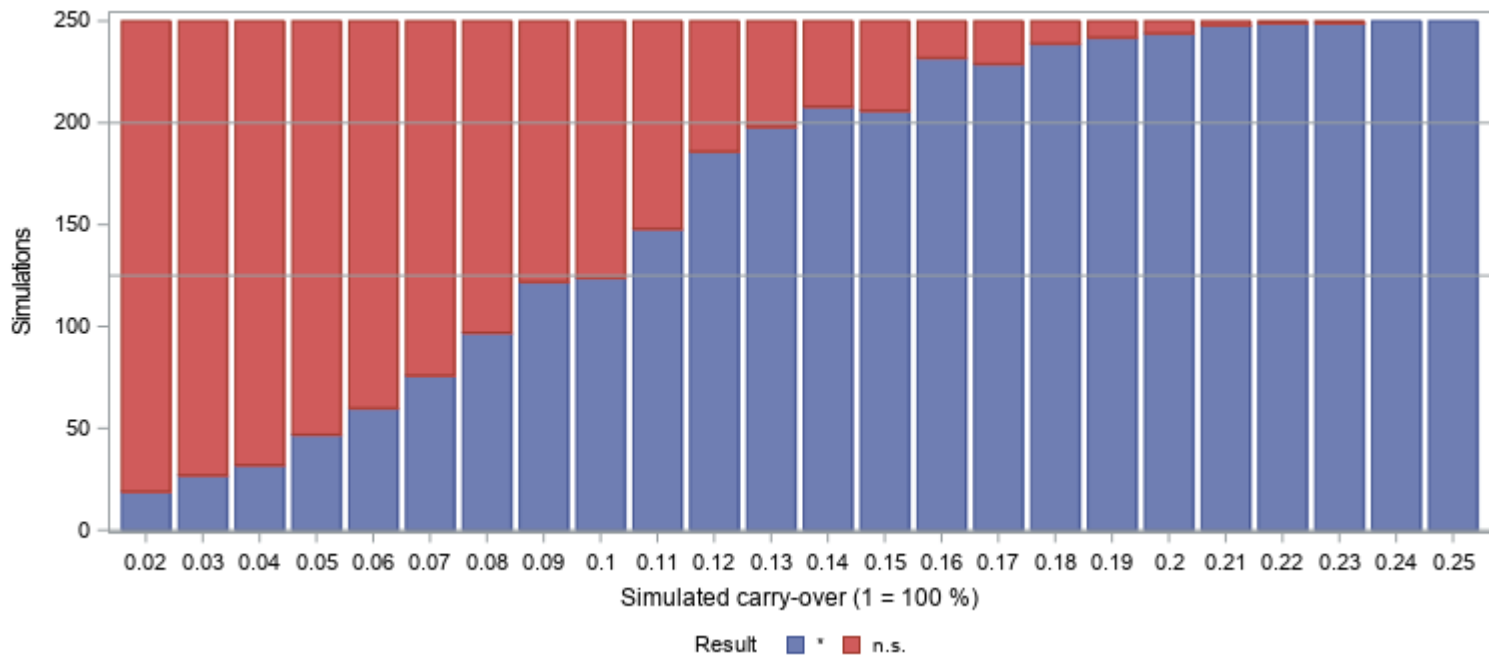
Results for farm test simulation with fat and protein contents (calculated carry-over)
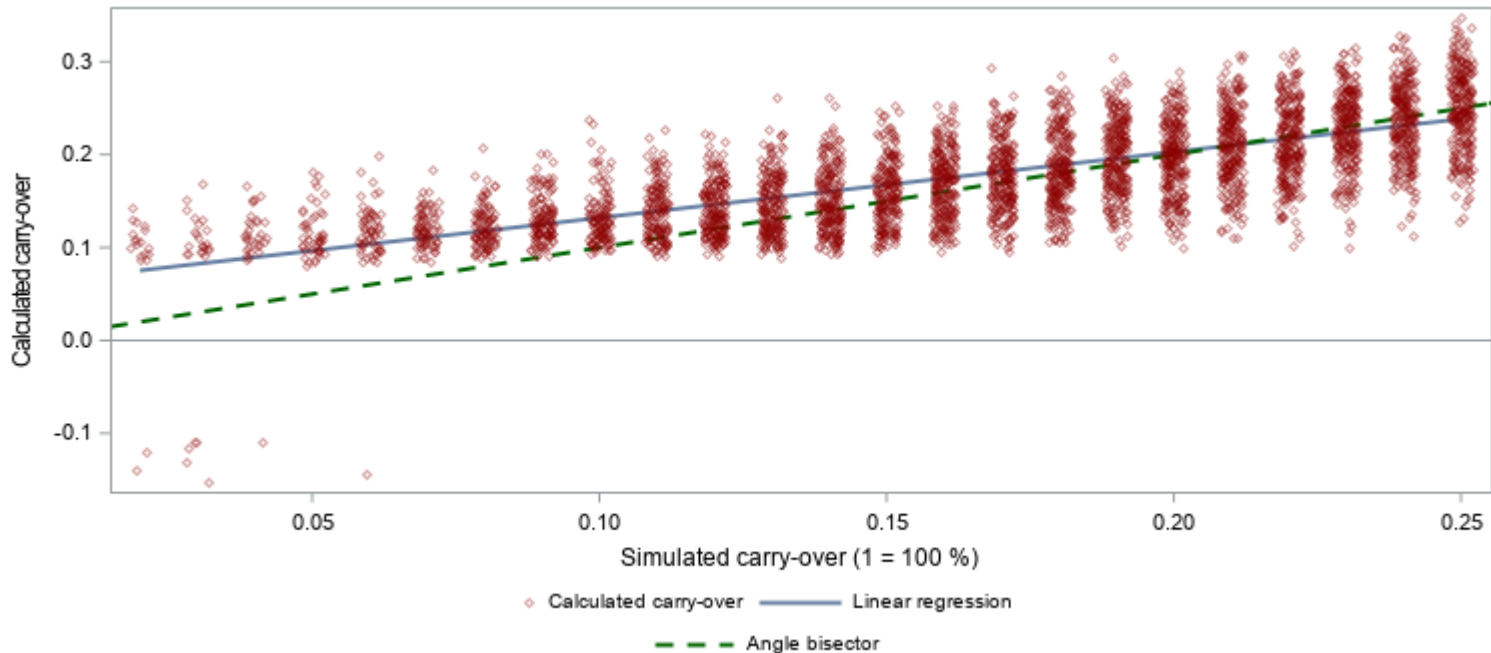
Results for farm test simulation with fat, protein and lactose contents (Power)

Results for farm test simulation with fat, protein and lactose contents (calculated carry-over)

Conclusion

- Running a statistical model with different farm test setups provided comprehensible results for carry-over estimates
- Impact of cow effect seems low and could maybe be neglected
- Power calculations show that adding more components will improve reliability of these estimates
- Methodically it is not the correct way to treat additional milk components as repetitions, but it gets the job done

THE GLOBAL STANDARD
FOR LIVESTOCK DATA

Outlook

- Check the simulation setup to find the source of the pattern in the calculated carry-over
- Switch to a multivariate statistical model that treats the different components correctly
- Begin to set up models for conventional milk meters in parlours, including the case that carry-over depends on milk yield
- Run the carry-over estimation on more data from previous ICAR farm tests

THE GLOBAL STANDARD
FOR LIVESTOCK DATA