

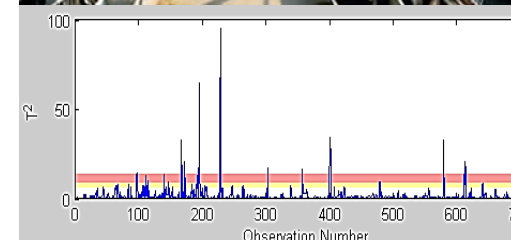


Principal component analysis for an early detection of mastitis and lameness

Bettina Miekley, Imke Traulsen, Joachim Krieter

Institute of Animal Breeding and Husbandry
Christian-Albrechts-University of Kiel

63rd Annual EAAP Meeting Bratislava, Slovakia
27th to 31st August 2012
Session 07





Introduction

- Mastitis and lameness most frequent and costly diseases
- Several studies of health monitoring (Cavero et al., 2008; Lukas et al., 2009; Pastell et al., 2009)
- Transfer problems into practice:
 - High error rates
 - High amount of false positive cows per day



Introduction

- Chemical and industrial process control:
 - Latent structure methods for fault detection
- Principal component analysis (PCA) combined with Hotelling's T^2 and residual (SPE) monitoring charts
 - Extension of univariate to multivariate monitoring of process control

Aim of this study: Applicability of PCA combined with control charts (T^2 and SPE charts) for an early detection of mastitis and lameness in dairy cows



Data

- Data:
 - Research farm Karkendamm, University of Kiel
 - Observation period: August 2008 until December 2010
 - 338 (mastitis) and 315 (lameness) cows in their first 200 days in milk (66,000 cow-days)
- Traits:
 - Milk yield [kg/milking]
 - Milk electrical conductivity [reference units/milking]
 - Daily pedometer activity
 - Feed intake [kg/day]
 - Number of feeding visits per day
 - Feeding time per day [minutes/day]
 - Mastitis and lameness treatments

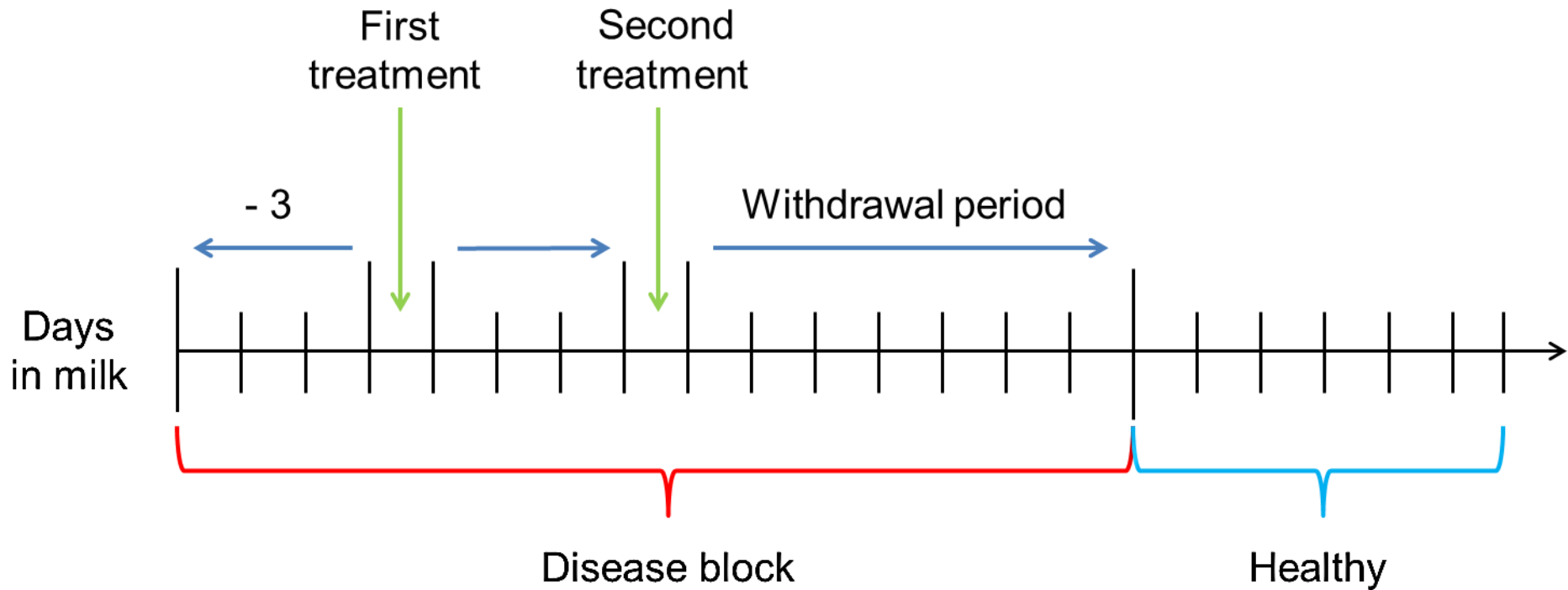


Definition of disease: Mastitis

- “Days of health” and “days of disease”
- Two mastitis definitions:
 1. **Mastitis + 3**: Day of treatment plus three days before
 2. **Mastitis + 4**: Day of treatment plus four days before
- Development of disease blocks= uninterrupted sequence of days of disease
- Accounting for early disease detection:
 - Analysis of days before the first treatment of each block



Definition of disease: Mastitis





Definition of disease: Lameness

- Three lameness definitions:
 1. **Lame + 3**: Day of treatment plus three days before
 2. **Lame + 5**: Day of treatment plus five days before
 3. **Lame + 7**: Day of treatment plus seven days before
- Development of blocks analogue to mastitis



Methods: PCA in general

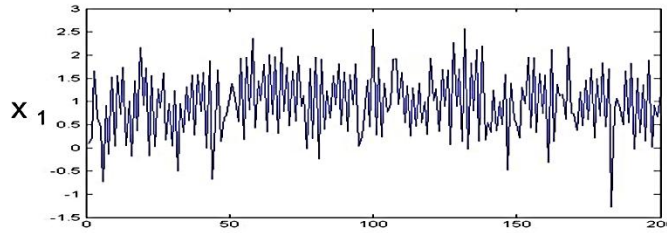
- **Aim:** Extraction of important information from (correlated) variables
 - New, uncorrelated and fewer variables:
Principal components (PC)
- Training dataset:
 - Establish PCA model: Only common cause of variation
 - 100 healthy cows over 200 days in milk
- Test datasets:
 - Healthy and ill cows



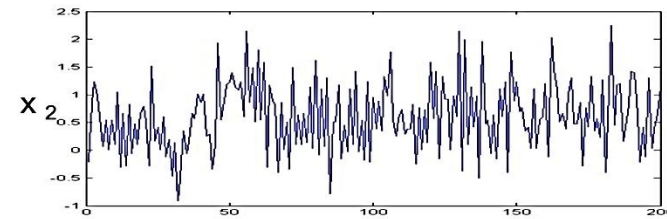
Methods: PCA in general

Original data

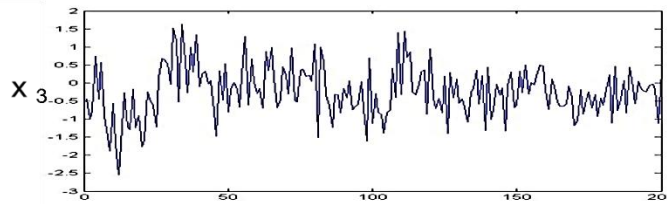
Milk yield



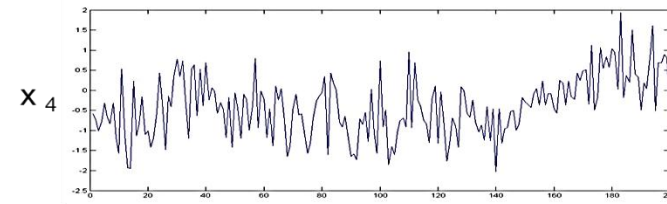
MEC¹



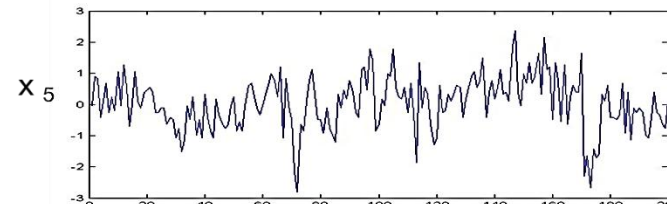
Feed intake



Feeding time



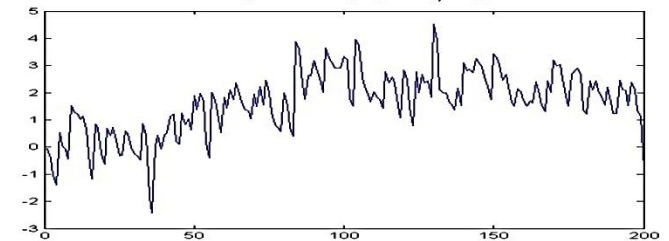
Number visits



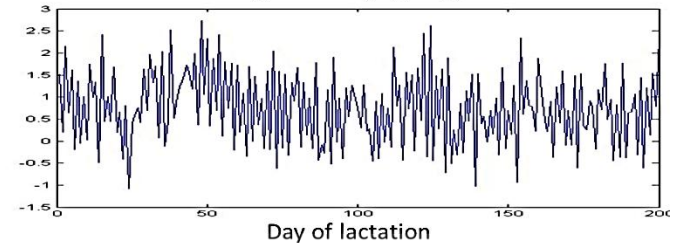
Day of lactation

Principle Components

t_1 from (x_3, x_4, x_5)



t_2 from (x_1, x_2)



¹MEC=Milk electrical conductivity



Methods: On-line monitoring

- New multivariate observations referenced against 'in-control' model

Off-line training

Historical data X



PCA

Variances,...

On-line monitoring

On-line samples X_{new}



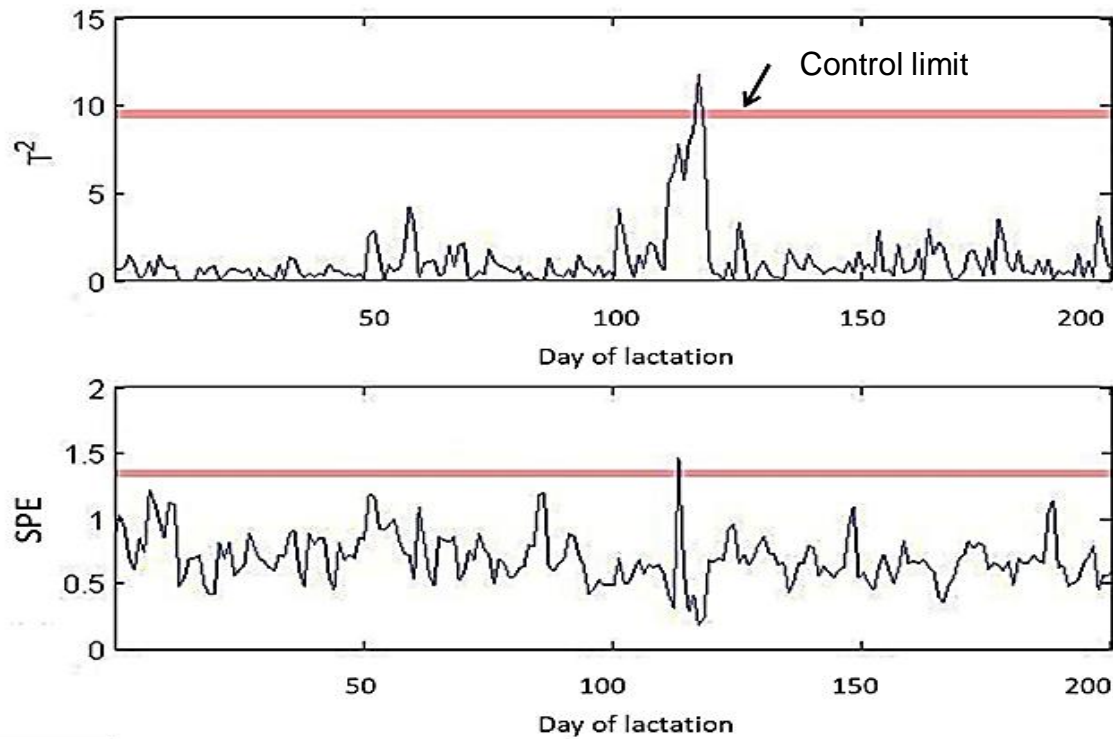
PC_{new}





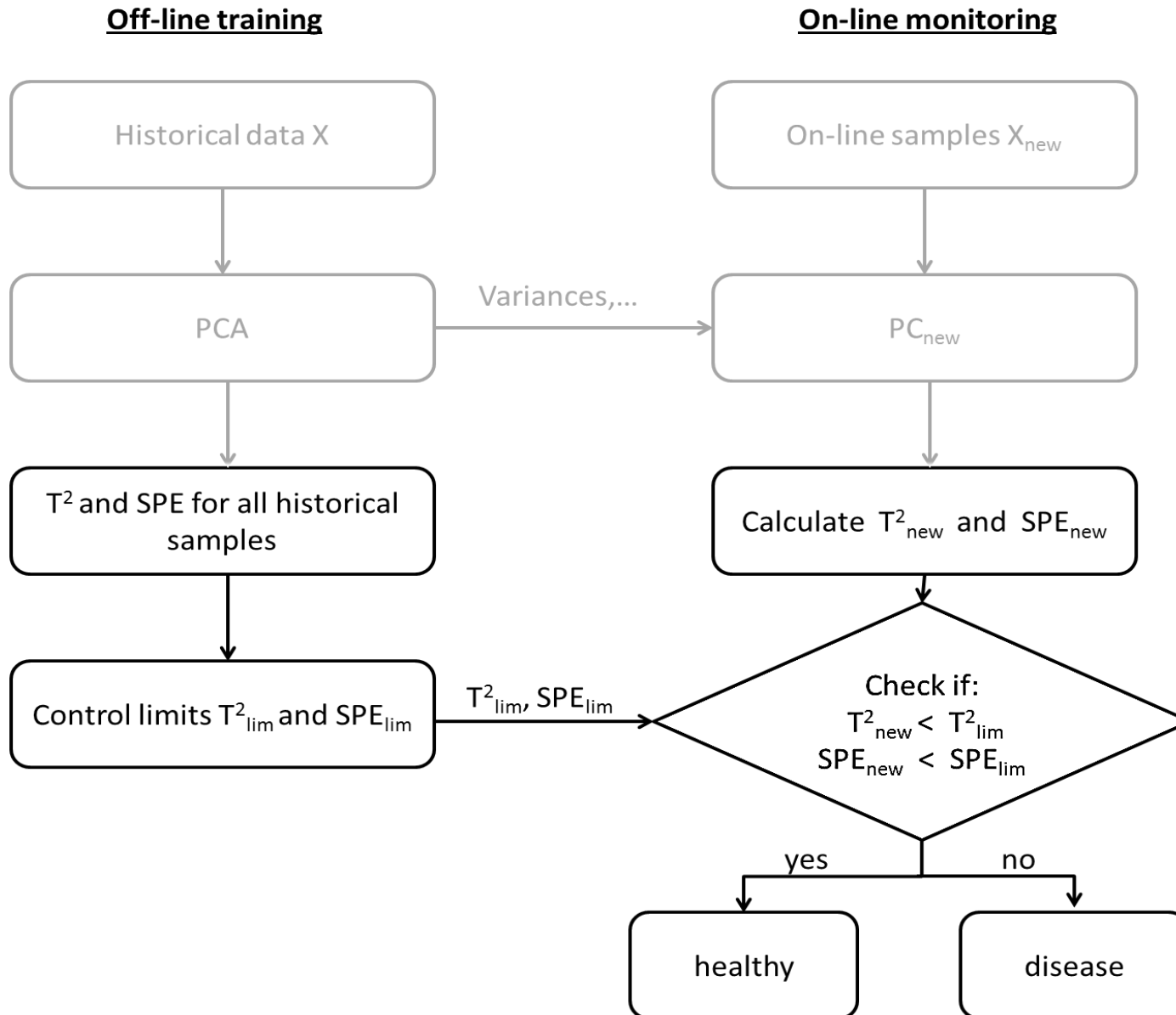
Methods: On-line monitoring

- Two complementary multivariate control charts for process monitoring:
 - Hoteling's T^2 chart
 - Squared prediction error (SPE) chart





Methods: On-line monitoring





Test procedure: Quality parameters

- **Sensitivity:** Percentage of correctly detected days of disease of all days of disease
- **Specificity:** Percentage of correctly detected days of health of all days of health
- **Error rate:** Percentage of days outside the disease periods of all the days where an alarm was produced
- **Block sensitivity:** Percentage of detected disease blocks within the days before a treatment

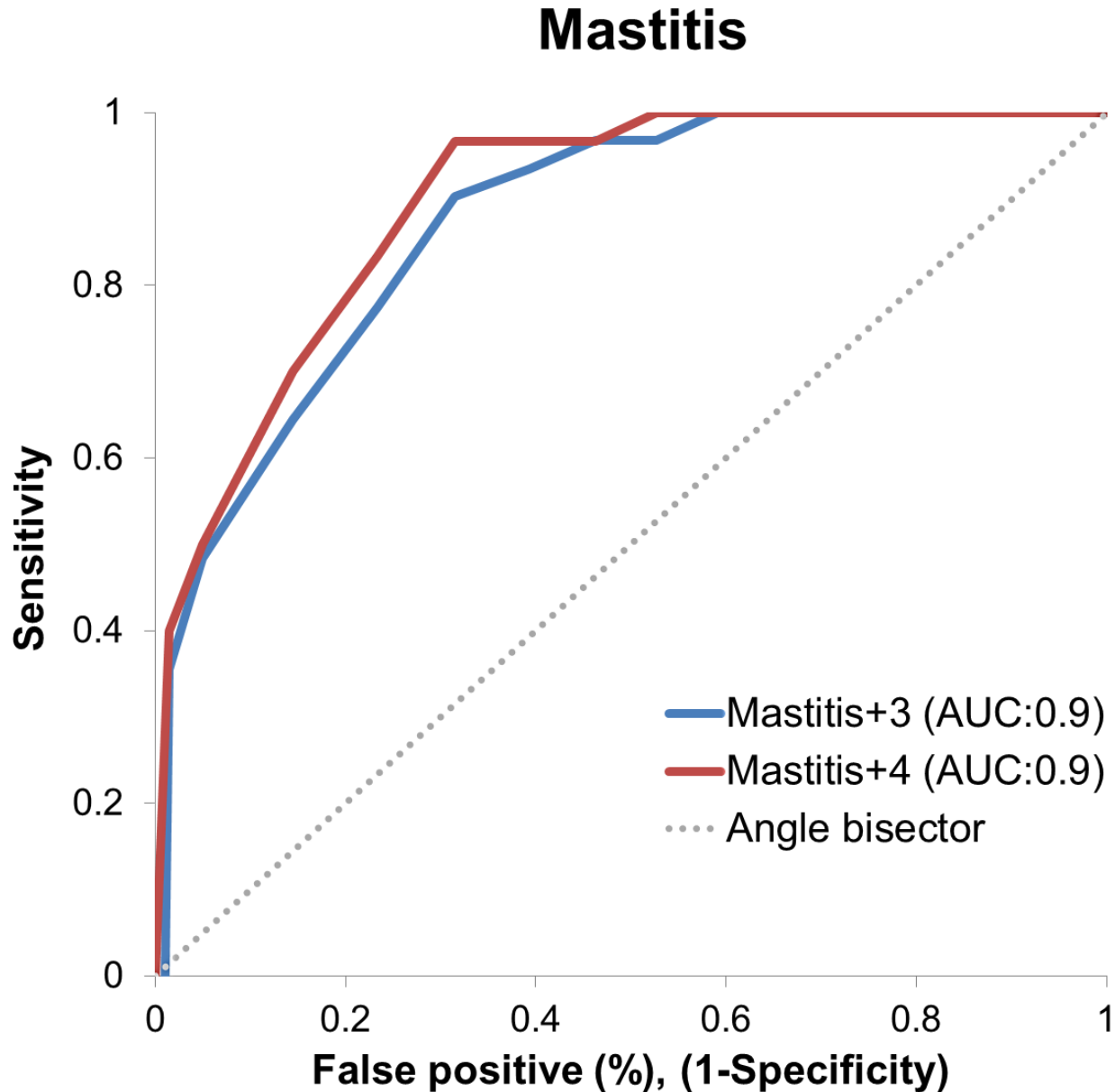


Test procedure: Quality parameters

- Tools for assessing accuracy of diagnostic predictions:
 - ROC (receiver operating characteristic) curves
 - X-axis: False positive fraction (1-specificity)
 - Y-axis: Sensitivity
 - Area under the curve (AUC)

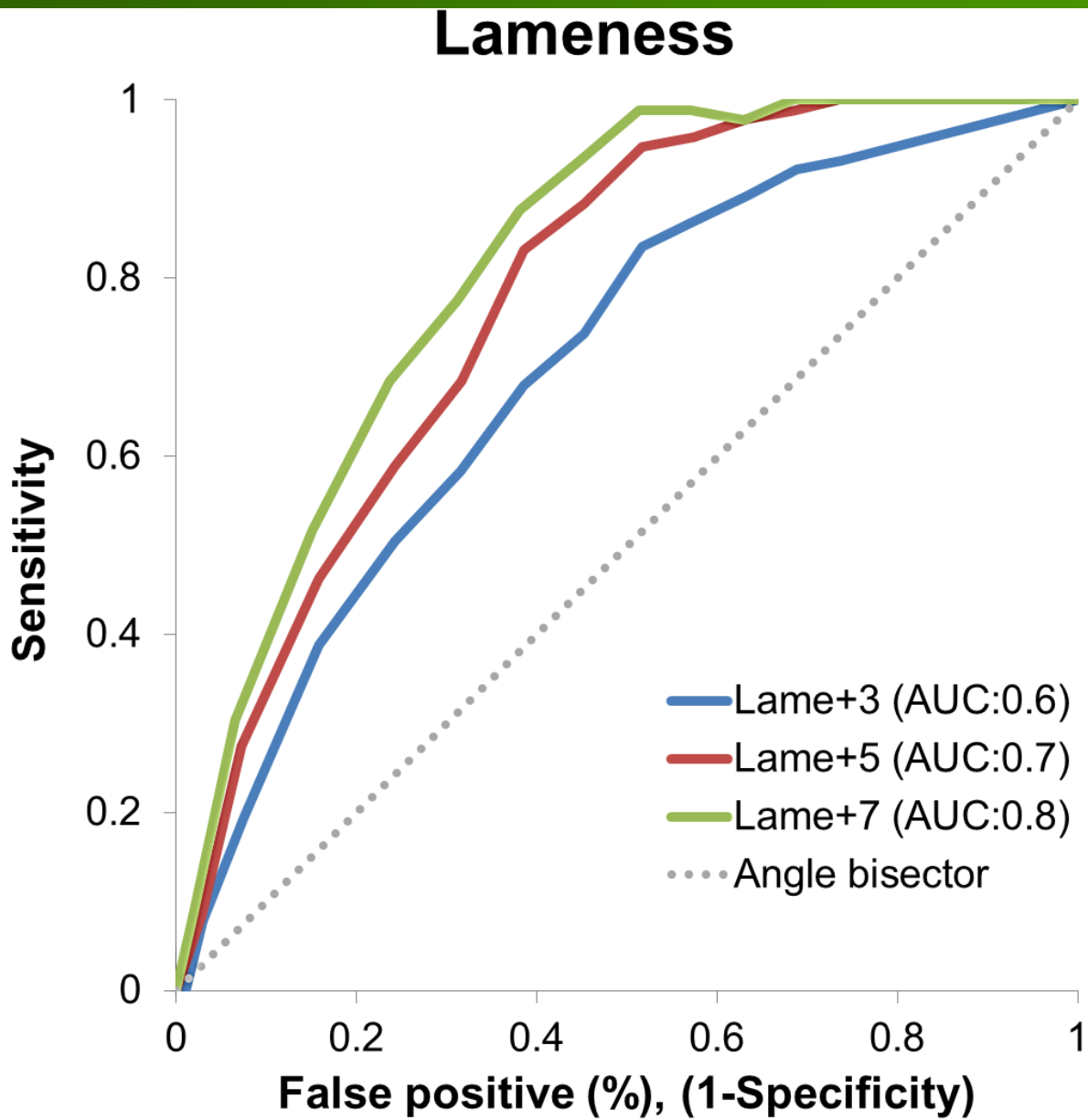


Results: ROC and AUC





Results: ROC and AUC





Results:Mastitis

Results of mastitis detection requiring a block sensitivity of least 70 %

	Block sensitivity [%]	Specificity [%]	Error rate [%]	TP cows/day	FP cows/day
Mastitis+3	77.4	76.7	98.9	0.2	15.2
Mastitis+4	83.3	76.7	98.8	0.2	15.0

Average herd size: 156 cows per day

FP(false positive): Cow incorrectly classified as ill

TP(true positive): Cow correctly classified ill

Mastitis+3: Treatment plus three days before

Mastitis+4: Treatment plus four days before



Results:Lameness

Results of lameness detection requiring a block sensitivity of least 70 %

	Block sensitivity [%]	Specificity [%]	Error rate [%]	TP cows/day	FP cows/day
Lame+3	73.8	54.8	89.2	1.3	12.3
Lame+5	83.2	61.4	88.5	1.3	9.9
Lame+7	87.8	61.9	87.8	1.3	9.3

Average herd size: 147 cows per day

FP(false positive): Cow incorrectly classified as ill

TP(true positive): Cow correctly classified ill

Lame+3: Treatment plus three days before

Lame+5: Treatment plus five days before

Lame+7: Treatment plus seven days before



Discussion

- Comparability to other studies difficult
 - Differences in disease definitions
 - Differences in block lengths (from 0 to 17 days)
 - Differences between sensors
- For 70 % block sensitivity:
 - Too high error rates
 - Too many FP cows/day



Discussion

- PCA combined with T^2 and SPE charts:
 - Capability of handling high-dimensional and correlated process variables
 - Easy to implement
 - Multivariate detection system
 - Discussions about PCA combined with other monitoring methods
- But:
 - No cow-individual analysis possible



Conclusion

- PCA possibility for disease detection
 - Without further performance improvement:
 - High error rates
 - Too many FP cows per day
- Impeding implementation into practice at present



**Thank you
for your attention!!!**





Appendix: Data overview I

Number of cows (healthy / ill) in the test and trainings dataset

Number of cows	Mastitis		Lameness	
	Training	Test	Training	Test
all	100	238	100	215
healthy	100	138	100	73
ill	-	100	-	142



Appendix: Data overview II

Trait	Mastitis				Lameness			
	Training		Test		Training		Test	
MY¹ (kg/milking)	18.2	(3.6)	18.4	(3.8)	18.0	(3.7)	18.0	(3.7)
MEC² (reference units/milking)	490.3	(32.0)	497.5	(34.9)	493.5	(35.6)	494.7	(36.0)
Daily activity (contacts/h)	32.1	(14.2)	32.8	(14.7)	32.7	(8.9)	30.9	(10.2)
Feed intake (kg/day)	39.9	(11.2)	39.5	(11.1)	40.6	(11.1)	39.0	(11.1)
Number of feeding visits per day	45.8	(13.7)	45.8	(14.1)	47.6	(14.0)	45.1	(13.8)
Feeding time (min/day)	177.3	(50.3)	176.3	(52.3)	181.0	(49.0)	176.5	(52.3)

¹MY=Milk yield, ²MEC = Milk electrical conductivity



Appendix: User interface

