Decision Support for Pneumonia Management in Pig Production

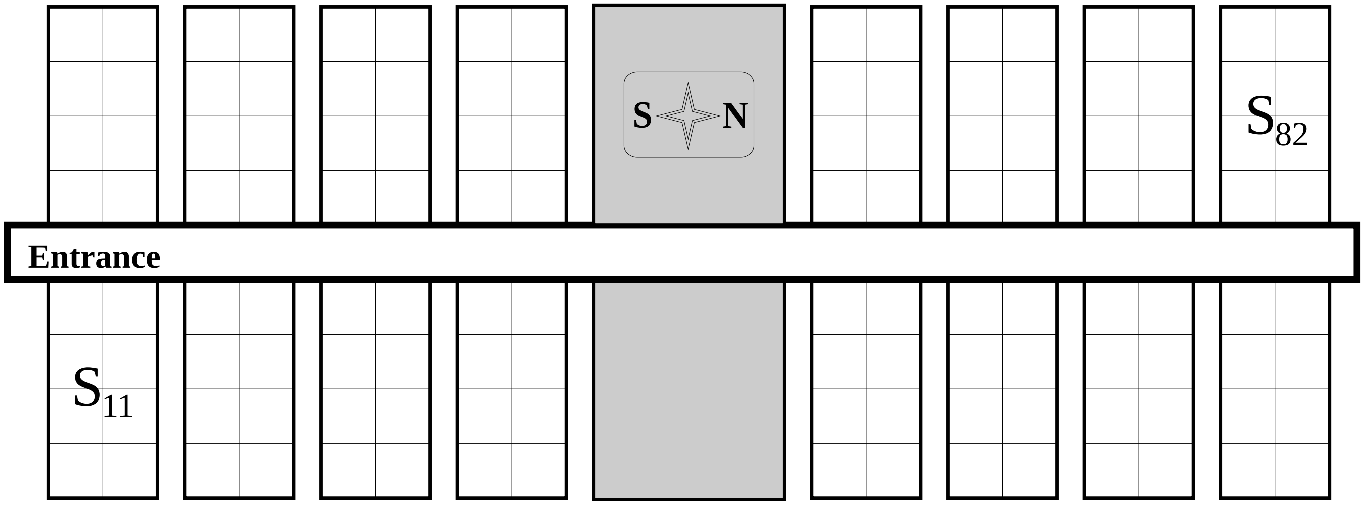
Ph.D. Stud. Michael Höhle
Department of Animal Science and Animal Health
Royal Veterinary and Agricultural University
Denmark

INTRODUCTION

- Aim of work: Exploit on-site treatment recordings to obtain insight and predict the spread of infectious diseases.
- Data provided by Danish National Comittee for Pig Production for boar test facility Bøgildgård.
- *Pneumonia* a term the workers use; covers visual detection of symptoms like panting, dry cough or inactivity. Causes are PRRS, mycoplasma hyopneumonia, pleuropneumonia, etc.

SPATIO-TEMPORAL ILLUSTRATION OF TREATMENTS

- Animation allows to review and analysis the events.



⋆ A green pen : Pen contains boars the specific day

⋆ A red pen : At least one boar in pen treated.

The Treatment Strategy

- Instance of sequential decision making under uncertainty.
  - ⋆ Daily decision; today's choice impacts tomorrow's state
  - ⋆ State only partially observable due to imprecise tests, etc.
- Two treatment regimes
  - ⋆ Individual treatments – injections with anti-biotics for 2-3 days
  - ⋆ Section treatment – anti-biotics in the water supply of the section.
- Idea for decision support: Predict occurrence of new cases → risk map.

Crafting the Risk Map

- Operate on daily level with section granularity.

$$Y_t^s = \begin{cases} 1 & \text{if one or more } \textit{new} \text{ infections appear in } s \text{ on day } t, \\ 0 & \text{if } \textit{no} \text{ new infections appear in } s \text{ on day } t. \end{cases}$$

- Use parametric model to compute prediction $\hat{Y}_{t+k}^s$.
- Colorize each section according to $\hat{Y}_{t+k}^s$ to obtain risk map for day $t+k$.
- Let map aid decisions, s.a. keeping a higher alert level, perform pre-emptive culling, apply water medication, etc.

Logistic Discrimination

- Assume prior probabilities π_0 and π_1 for the two states of Y_t^s and misclassification losses L_{01} and L_{10}.

- Compute posterior $p(Y_t^s|x_t)$ by some parametric model using observed covariates x_t

- Minimum loss Bayes rule based on posterior

$$c(x_t) = \begin{cases} 1 & \text{if } p(Y_t^s = 1|x_t) > L_{01}/(L_{01} + L_{10}) \\ 0 & \text{otherwise} \end{cases} .$$

Classifier Training and Evaluation

- Assume 1:1 correspondence between treatment and disease and split dataset into training and validation sets.

- Use confusion matrix to calculate misclassification rates.

$c(x_t)\backslash Y_t$	1	0
1	n_{11}	n_{01}
0	n_{10}	n_{00}

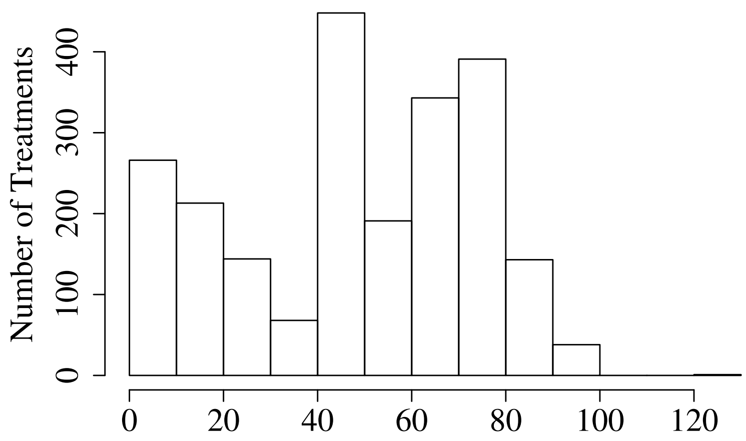
$$\begin{aligned} Se &= \frac{n_{11}}{(n_{11}+n_{10})} \\ Sp &= \frac{n_{00}}{(n_{01}+n_{00})} \end{aligned}$$

- Evaluation metric – expected cost per case

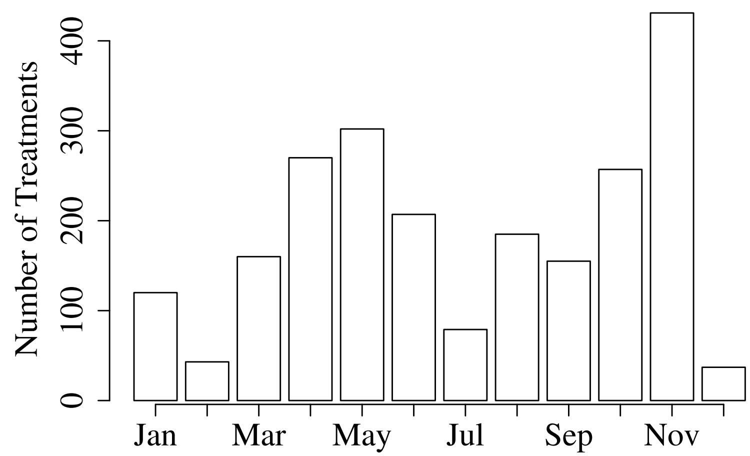
$$p(Y_t^s = 1)(1 - Se)L_{10} + (1 - p(Y_t^s = 1))(1 - Sp)L_{01}.$$

GENERALIZED AUTOREGRESSIVE MODEL (1)

- Time series model with discrete response conditioned on past.
- Histogram plots reveal effect of average age of boars in section, number of boars, and time of the year.



Age



Time of year

GENERALIZED AUTOREGRESSIVE MODEL (2)

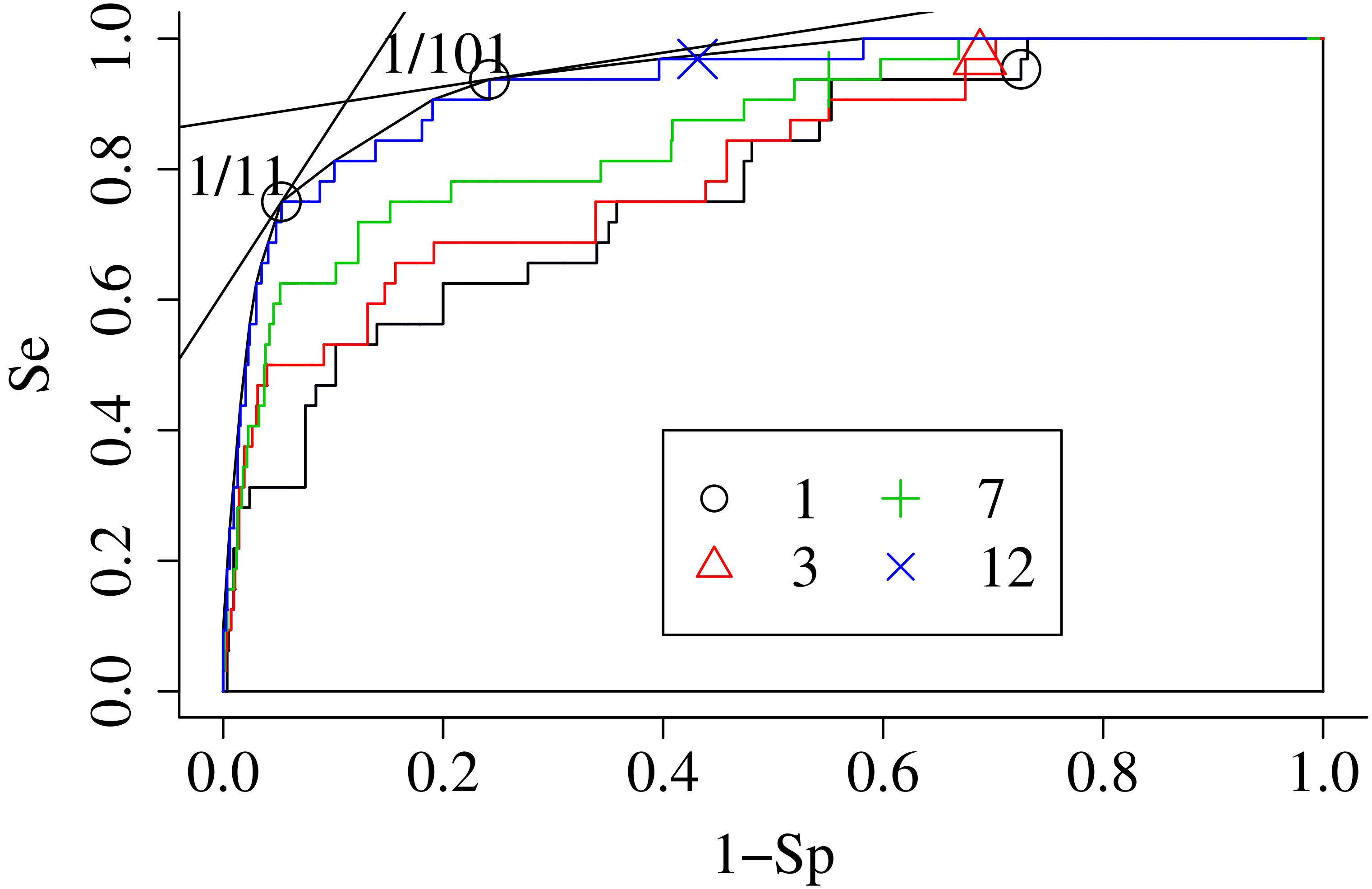
- Disease spread is modeled by including state of nearest compass direction neighbors.

- Resulting logistic link $GArM(l)$ model

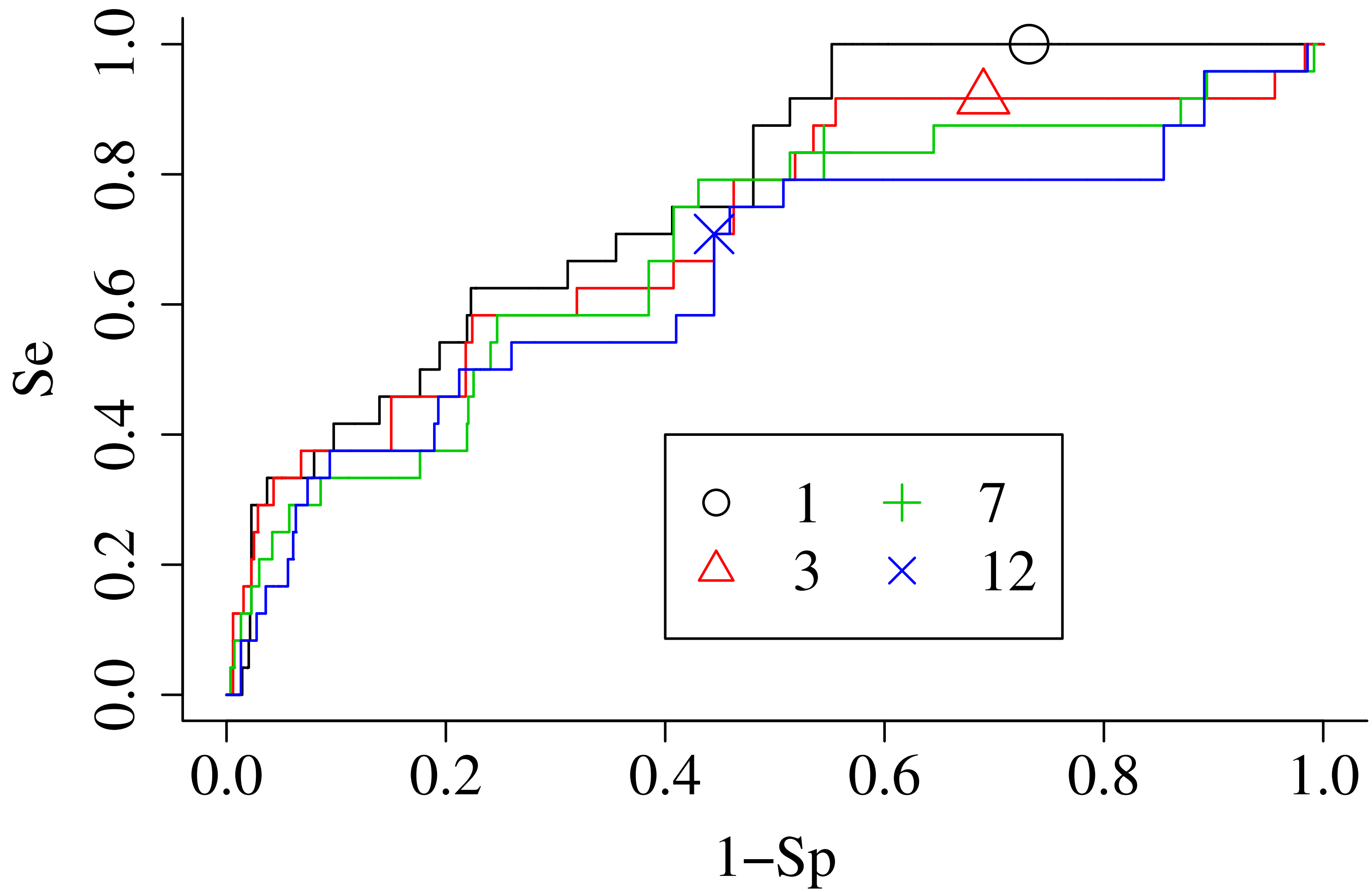
$$\text{logit}(\mu_t^s) \;=\; x_{\text{no}}^s\gamma_1 + x_{\text{age}}^s\gamma_2 + \gamma_3^{\text{season}} + \beta_0 + \sum_{s' \in N_4^*(s)} \sum_{i=1}^{l} \beta_i^{s'} y_{t-i}^{s'},$$

- Finding appropriate l is a model selection issue.

ROC CURVE FOR S_{11} – TRAINING SET



ROC CURVE FOR S_{11} – VALIDATION SET



Conclusion and Discussion

- Decision aid by predicting location of new cases.

- Black-box model is not able to find many systematic patterns in Bøgildgård data. White-box approach s.a. SIR-model might facilitate data better.

- Treatments yield only partial information on disease state. Explicitly modeling this fact might be beneficial, but hard to quantify.

- Retrospective analysis immediately useful. Prediction system a step towards the goal of decision support systems in health management.