

# Statistical evaluation of artificial intelligence systems

## Project II

January 2020

### Phosphate in soil and the effect on barley production

**Background** Phosphorous (P) in form of phosphate is an essential nutrient for plants. It stimulates plant growth, and lack of phosphorus in soil affects harvest yields negatively.

However, most of the phosphorous in the soil is "bound" and cannot be utilized by the plants. It is thus of crucial interest for farmers to get a good measure for the amount of bioavailable phosphorous.

**Experiment** Soil samples were obtained from nine fields in Denmark and Norway. Each field was divided into a number of plots, and thereafter barley was sown on the plots. The soil samples were analysed for bioavailable phosphorous using two methods, "Olsen-P" (the traditional method) and "DGT" (a newer, more expensive method).

**Data** Data is stored in `fosfor_data.RData`. There are four variables:

- `location`: ID for field location, (001 - 011)
- `yield`: Harvest yield of barley in hkg/ha. (hkg: 100 kg)
- `DGT`: DGT measurement in  $\mu g/L$ .
- `olsenP`: Olsen P measurement in mg/100g.

**Aim of project** Analyse the data and make relevant conclusions and recommendations for farmers, focusing on:

- choosing between Olsen P and DGT measurements.
- does the amount of bioavailable phosphorous influence the harvest yield?

**Remarks** You may consider using the *Michaelis-Menten model* (you do not have to use this model), which is often used for this kind of data. The Michaelis-Menten model assumes

$$y \approx \frac{\alpha \cdot x}{\beta + x}, \quad x > 0$$

where  $y$  is modeled as a function of  $x$ .  $\alpha$  and  $\beta$  are parameters in the model;  $\alpha$  is the value of  $y$  as  $x \rightarrow \infty$ .

The Michaelies-Menten model is an example of a non-linear regression, which can be fitted using `nls` in R. See supplementary material for an example of non-linear regression in R. Additional info can be found in Crawley, pp. 142-146.

## Supplementary material

The `nls` function can be used to fit non-linear regression in R with a least-squares criterion (ie. normally distributed noise). Example:

```
phos.model <- nls(yield ~ alfa * DGT/(beta + DGT) , data = Phosphorous,  
start = list(alfa = 90 , beta = 1))
```