Using Data from Literature for Fuzzy Rule based Modelling of Nitrate Leaching

Sven van der Heijden, Uwe Haberlandt

Aims

The purpose of this study was to examine to what extent it is possible to calculate mean annual values of nitrate leaching based solely on data collected from literature (“collective knowledge”). The focus was on nitrate leached from agriculturally used areas under a diverse set of conditions. For this task a simple fuzzy rule based model was employed.

Data

A total of 222 data records were collected from published studies. Apart from the dependent variable “nitrate leaching”, these records comprise mean annual values of precipitation, fertilization and percolation rate as well as information on soil type and cultivated crops. For the latter two, aggregated indices were derived from more detailed data.

Only one study provided measured data from lysimeter analysis, the others provided simulation data from different models.

Results

- Two fuzzy models with 4 and 5 input variables respectively and 30 rules each have been tested and compared to a multiple linear regression with 4 variables.

Tab. 2: Results of 100-fold split sampling validation for nitrogen leaching simulation

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Average (kg ha(^{-1}) yr(^{-1}))</th>
<th>RMSE (kg ha(^{-1}) yr(^{-1}))</th>
<th>Correlation (r)</th>
<th>Training</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>TRN</td>
<td>VAL</td>
<td>TRN</td>
<td>VAL</td>
<td>TRN</td>
</tr>
<tr>
<td>1</td>
<td>Fuzzy Rules 4 Variables: SI, P, F, CI</td>
<td>42.7</td>
<td>43.3</td>
<td>10.5</td>
<td>21.6</td>
<td>0.95</td>
</tr>
<tr>
<td>2</td>
<td>Fuzzy Rules 5 Variables: SI, P, F, CI</td>
<td>42.7</td>
<td>43.3</td>
<td>11.5</td>
<td>21.2</td>
<td>0.95</td>
</tr>
<tr>
<td>3</td>
<td>Multiple Regression 4 variables: SI, P, F, CI</td>
<td>42.7</td>
<td>43.4</td>
<td>23.7</td>
<td>25.1</td>
<td>0.79</td>
</tr>
</tbody>
</table>

- The fuzzy models have a satisfying performance and give better results than the multiple linear regression.
- The fuzzy models explain almost 70% of the variance of the validation data.

Fig. 3: Relationship between “observed” and simulated nitrogen leaching using fuzzy model with 4 variables averaged over 100 split samples with 150 training records and 72 validation records each

- A quantitative prognosis of nitrate leaching based only on literature data is possible.
- Fuzzy rule based modelling is a suitable tool to store and interpolate “collective knowledge”.
- The tool should be well suited for incorporation into a DSS and for large scales (e.g. EU-wide, this study already included data from Sweden, Denmark, Germany, and France)

Methods

- Fuzzy rule based modelling

  \[ F (x_i: A_i) \land ... \land F (x_n: A_n) \rightarrow F(y: B) \]

  \[ \mu_i = \text{Fuzzy number } \rightarrow \mu_M \]

  \[ \mu_r = \text{Fuzzy response } \rightarrow \mu_y \]

  - Input variable \( x \rightarrow y \)
  - Dependent variable \( y \)
  - Membership function \( \mu_i \)
  - Fuzzy response \( \mu_y \)

- Random division of input data into training and validation sets (150:72 records)
- Generation of an optimized rule system from the training data set using simulated annealing (discrete, non-linear, numerical optimization) with the objective function:

  \[ O = \sum_{i=0}^{n} (y_i - \hat{y}_i)^2 \rightarrow \min \]

  \( y_i \) = Input value of nitrogen leaching

- To reduce uncertainty the split sampling is repeated 100 times

Outcome

- Requirement for such applications is a careful preparation of publications with detailed information about data and results.

Outlook

- With smaller time steps (seasonal or monthly) results should improve.
- Problem: data availability
- Although fuzzy modelling allows for missing data in the variables, the inclusion of such a variable did not improve the result.
- Probably also a problem of too large time steps, since the percolation rate should have a major impact on nitrate leaching.

Literature