Towards selective hoeing depending on evaporation from the soil

Christoph Manss¹, Kai von Szadkowski², Janis Bald³, David Richard³, Christian Scholz³, Daniel König³, Matthias Igelbrink³ and Arno Ruckelshausen³

Abstract: This paper presents how to generate an artificial dataset to test different hoeing rules. Therefore, images that have been obtained on two days of a field trial are analysed to infer weed and crop sizes. Then, weather data from 2021 and 2022 is gathered from open-source data for 100 synthetically generated fields. The generated dataset is then used to test hoeing rules that are conditioned to keep as much moisture in the soil as possible. The analysis with these hoeing rules indicates that much less hoeing would be applied if the proposed hoeing rules are used.

Keywords: weed control, data set generation, expert systems

1 Introduction

Water management is becoming increasingly relevant for sustainable agriculture. The world is facing more extreme weather events and their duration and succession becomes more unpredictable [SB21]. Consequently, farmers must change their usual practices to adapt to the new climate. One topic that will play a vital role in agriculture is water management [BBŠ22]. Heavy rain events need to be diverted and the field must be prepared for dry weather periods. It may thus be beneficial to implement rules for soil tillage that take its impact on water management of the soil into account.

Usually, soil tillage focuses on preparing the ground for seeding and weed management [Ge22]. Nowadays, zero tillage is considered as it might have better long-term effects on the soil regarding water management, e.g. evaporation [Bh17; Li22; Na21]. In the recent years, the effects of weed control in connection with tillage and zero tillage have been studied [Bu22; De20]. For tillage, the soil is cut deeply or turned over, but for weed control, only the surface of the soil (2.5 cm to 10 cm) is tilled to destroy unwanted weeds. In [BBŠ22], the authors show that weed control can reduce the soil moisture significantly. Depending on the current soil condition, this can lead to weaker plant growth. Yet, weed...
control is also vital as the weeds compete with the crops among other things for water in
the soil.

Thus, in this paper, rules for selective hoeing in maize are evaluated with the constraint to
keep as much water in the soil as possible. Consequently, this approach looks not only at
the size of the weeds, but also at additional circumstances in the field such as previous
weather conditions, weather forecast, soil attributes, and the development stage of the
crops. To evaluate these rules, a data set based on weather data of Lower Saxony (DE)
and soil statistics is generated to evaluate rules of selective hoeing mechanisms.

2 Generation of a synthetic data set

<table>
<thead>
<tr>
<th>Feature</th>
<th>Unit</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPS coordinate (latitude, longitude, height)</td>
<td>(°, °, m)</td>
<td>Random points to get weather data</td>
</tr>
<tr>
<td>Soil types</td>
<td>-</td>
<td>See Tab. 2</td>
</tr>
<tr>
<td>Seed date</td>
<td></td>
<td>Between mid-April and mid-Mai</td>
</tr>
<tr>
<td>Hoeing date</td>
<td></td>
<td>4 weeks after seeding</td>
</tr>
<tr>
<td>Max temperature when hoeing</td>
<td>[°C]</td>
<td></td>
</tr>
<tr>
<td>Max temperature 3 days after hoeing</td>
<td>[°C]</td>
<td></td>
</tr>
<tr>
<td>Sun hours {on the day of/two days after}</td>
<td>[h]</td>
<td></td>
</tr>
<tr>
<td>hoeing date</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics on maize size</td>
<td>[m²]</td>
<td>Inferred from obtained data (mean and</td>
</tr>
<tr>
<td>Statistics on weed coverage</td>
<td>[m²]</td>
<td>standard deviation)</td>
</tr>
<tr>
<td>Field size</td>
<td>[m²]</td>
<td></td>
</tr>
<tr>
<td>Rain 3 days {before/after} hoeing date</td>
<td>[mm]</td>
<td></td>
</tr>
</tbody>
</table>

Tab. 1: Descriptive summary of the synthetic data set

To evaluate the decision system, a synthetic data set is generated. The data set is arranged
in a tabular fashion and comprises the features as shown in Tab. 1. The assumptions and
how the data has been generated are described in the following sections.
2.1 Assumptions and simplifications for the synthetic data set

To generate the synthetic data set the following is assumed:

- Maize is planted in between mid-April and mid-May. Then, the first hoeing application follows four weeks after seeding respectively, thus aligning with BBCH growth stages 12 to 16. Often a second hoeing application is applied six weeks after seeding, but this is not considered in this work.

- Fields in Lower Saxony were considered, where an agricultural holding has on average 73 ha of land [La22, Ni21]. Furthermore, it is assumed that this land is divided into 14 fields, which yields a field size of 5 ha. It is admitted that field sizes can vary according to a Gaussian distribution with a standard deviation of 0.3 ha.

- We assume only one soil type per field. The soil type is selected by the likelihoods in Tab. 2.

<table>
<thead>
<tr>
<th>Sand</th>
<th>Clayey sand</th>
<th>Loamy sand</th>
<th>Strong sandy clay</th>
<th>Sandy clay</th>
<th>Clay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sand</td>
<td>0.2</td>
<td>0.15</td>
<td>0.25</td>
<td>0.2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Tab. 2: The assumed likelihoods of the soil types, which might appear on the farmland

2.2 Plant Data

For the plant data, images of a maize field that has been recorded on two days in July 2021 were utilized. The system for the recording is presented in [Ko22]. The images have been manually labelled with the classes maize and weed (see Fig. 1). Crop rows, number of removable weeds, and coverage of the maize plants are inferred from the resulting box annotations as explained in the following. For the identification of the crop line, a regression is calculated based on the centre positions of the maize boxes. This regression

![Detected maize plants and weeds (left) and estimated cropline (right). According to the width of the maize, a margin is evaluated to identify in-row and off-row weeds. The off-row weeds can be removed with the hoe](image-url)
represents the crop line, as shown in Fig. 1. To get the number of weeds that can be removed by selective hoeing, a margin from the crop line is calculated likewise. Subsequently, weeds that are not within the margins are removed. This is computed for all labelled images. Then, the resulting data is used to predict the weed sizes per individual day by employing an exponential function

\[ f(t) = a \exp(bt) + c, \]  

(1)

where \(a, b, c \in \mathbb{R}\) are the parameters to be learned and \(t \in \mathbb{N}\) are the days. The parameters are inferred using a nonlinear least-squares optimization. To extract the data for both days, we make use of the Python package Plantcv [Fa15]. The pre-processing for the size determination of each plant is shown in Fig. 2. In this process, the image is first white balanced and converted into the LAB colour channel, where we only use channel A. Every value in the A channel is then set to zero if it is below a threshold and set to one if otherwise. This creates a binary mask. Then objects in the image are identified and filled with zeros, if they are below a certain size. Lastly, the remaining objects are dilated to reduce the gaps in the remaining objects. After the pre-processing, the bounding boxes are applied on the binary mask and the pixels equal to one in each bounding box are counted. The geometric size of the pixels is known such that a size in square metres can be computed.

For maize, another function \(g(t)\), which uses the same model as in (1), is fitted to the size of maize. The results of both regressions are shown in Fig. 3 on the left together with the densities of the weed and maize sizes on the right. This regression is then used to extrapolate weed growth off row and maize growth for the synthetic data.
Towards selective hoing depending on evaporation from the soil

2.3 Weather data

Now that the time for seeding and land tillage are determined, the weather data can be gathered. Therefore, the GPS position of each field instance is used to gather the actual weather of this position for the hoing date. For this, we utilize the python package Meteostat [Ch22], which accesses multiple open data sources to get weather data.

3 Derivation of an expert system

In this section, decision rules are presented that decide according to the data if hoing should be applied or not. In this paper, four rule systems are considered: conventional hoing, selective hoing, and two evaporation-optimized rule sets for hoing. The rules are based on empirical studies as well as recommendations of agricultural societies. 

Conventional hoing is basically the current standard in land tillage; the entire field will be tilled.

For selective hoing, only the size of the maize and the weeds are considered. The following rules are considered:

- The BBCH of the maize must be within 12 and 16. This is measured by the size of the maize $s_{maize} \in \mathbb{R}^+$ in [m$^2$]. Therefore, $0.01 \text{ m}^2 \leq s_{maize} \leq 0.3 \text{ m}^2$ for hoing.

- The weed overage $c_{weed} \in \mathbb{R}^+$ in [m$^2$] must be large enough. Thus, $c_{weed} > 0.002 \text{ m}^2$ to activate the hoe. That means that if all weeds are accumulated in a single patch, the patch has an edge length of $4.5 \text{ cm} \times 4.5 \text{ cm}$.

If all the rules are fulfilled, the hoing is activated.

Evaporation-optimized hoing considers rules that are chosen to maximize the water content in the soil or to minimize the evaporation according to the current literature. Soil evaporation in dependence of tillage has been looked at in [WB71] and for weed control in eucalyptus plantations in [De20]. However, for this scenario, the literature is sparse.
Thus, the chosen values are based on educated guesses: sunny weather dries out the ground [BK15]. Loamier soils can hold the water better and can therefore be hoed even if there is less weed. If the weather is rainy, hoeing is not so harmful regarding the moisture in the soil.

This paper considers two rules for evaporation-optimized hoeing: $E\text{-hoeing}1$ and $E\text{-hoeing}2$. For $E\text{-hoeing}1$ we consider the same rules as for selective hoeing and the sun. Thus:

- The sun should not shine longer than 8h per day $s_t > 8$ h/d

For $E\text{-hoeing}2$, we consider a more complex rule-system, which takes the type of soil into account together with the weather conditions. The rules for E-hoeing2 are the following:

- The weed coverage $c_{\text{weed}}$ is put into relation to the soil type. Here we assume that clay can hold water much better compared to sand. Consequently, clay will be hoed already if $c_{\text{weed}}$ is small, and sand will be hoed if $c_{\text{weed}}$ is large. For this rule, we chose a threshold $c_t > 0$ in $[m^2]$ that is linearly increasing in equidistant steps with the soil type. For clay $c_{t, \text{clay}} = 0.002$ m$^2$ and for sand $c_{t, \text{sand}} = 0.01$ m$^2$.

- Rainy days can influence how effective the hoeing can be. If it is rainy after hoeing, removed weeds might regrow. Thus, normally it is suggested that hoeing is applied before a period of dry weather. Here, we choose a different rule as we want to prepare the soil for the rain. Therefore, we set a threshold for the amount of rain within the next three days $r_t > 7$ mm to activate the hoeing.

- Lastly, if the weather is sunny and hot, it is considered not to apply the hoeing because hoeing might increase the ground surface area such that more water

---

![Fig. 4: After hoeing, this figure shows on the left y-axis the hoed ground and on the right y-axis the weed coverage, dependent on the threshold parameter for the weed size $c_{\text{weed}}$. Additionally, some of the soil dependent thresholds are annotated.](image-url)
Towards selective hoeing depending on evaporation from the soil

Evaporates. Therefore, the thresholds \( s_t < 8 \text{ h/d} \) and \( t_t < 27 ^\circ \text{C} \) must be fulfilled for hoeing.

Hoeing is applied only if all these conditions are met. This paper only evaluates these rules and does not put them into practice. However, the interested reader might look into [Ni21_2], where the authors present how to implement such an inference system.

4 Evaluation

All systems are compared regarding the removed weeds and the hoed ground. Furthermore, statistics on soil evaporation are utilized to estimate the water loss in the ground in each scenario. Fig. 4 shows the influence of \( c_{\text{weed}} \) on the hoed ground and the removed weeds. The crossing point of hoed ground and weed coverage after hoeing seems like a sweet spot for the selective hoeing.

Each rule set is evaluated for a relatively wet year (2021) and a relatively dry year (2022). Thus, two datasets were generated. In Tab. 3, the hoed ground, averaged over all fields, is presented for both years. Clearly, the hoeing optimized for less evaporation hoes less soil by design. Compared with conventional hoeing, selective hoeing reduces the hoed ground by 40% and the evaporation hoeing reduces the hoed ground by almost 92 to 99%.

<table>
<thead>
<tr>
<th>Weeding rule</th>
<th>Weeded ground [m²]</th>
<th>Weeded ground relative to no-rules hoeing [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2021</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No-rules hoeing</td>
<td>50939</td>
<td></td>
</tr>
<tr>
<td>Selective hoeing</td>
<td>39662</td>
<td>77.86</td>
</tr>
<tr>
<td>E-hoeing1</td>
<td>3988</td>
<td>7.83</td>
</tr>
<tr>
<td>E-hoeing2</td>
<td>682</td>
<td>1.34</td>
</tr>
<tr>
<td>2022</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No-rules hoeing</td>
<td>50908</td>
<td></td>
</tr>
<tr>
<td>Selective hoeing</td>
<td>39610</td>
<td>77.81</td>
</tr>
<tr>
<td>E-hoeing1</td>
<td>2215</td>
<td>4.35</td>
</tr>
<tr>
<td>E-hoeing2</td>
<td>643</td>
<td>1.26</td>
</tr>
</tbody>
</table>

Tab. 3: Averaged results from the hoeing rules for the considered parameters
If the years are compared with each other, the hoeing rules for evaporation hoe a little more ground, if the weather conditions are more wet as in 2021 compared with dry weather conditions in 2022.

Next, we analyse the dependency of soil for each hoeing rule. Figure 5 displays the hoed ground depending on the soil for the weather in 2021 and Figure 6 for the weather in 2022. Comparing those figures, the E-hoeing2 is almost only applied if the soil is clay. For the other soils the weeds are not large enough. This can be seen as an artifact from the synthetic data set. The detected weed sizes are simply too small on those dates such that the threshold of weeds per image per soil type is too large. The E-hoeing1, on the other side, is more weather-dependent. In the wet year, it is applied more often than in the dry year. The reason why E-hoeing1 is more often applied for clayey soil is however a coincidence, as in the synthetic data soil and weather conditions are not linked.
5 Conclusion

This paper showed how to generate an artificial data set to test hoeing conditions for future smart agriculture applications in maize. For the data set, images from a field experiment were analysed to determine statistics of the maize and weed coverage. According to these statistics, multiple fields are generated with different soil properties and weather conditions on the date of hoeing. Given the simplicity of this artificial data set, first analyses can be applied regarding the conditions for hoeing. As test conditions, three expert systems are proposed that decide if the ground should be hoed or not based on the constraint to keep as much moisture in the soil as possible. The presented analysis shows that these systems hoe much less soil. The systems also have been tested against two datasets with dry and wet weather conditions. For the wet weather, the hoing was activated much more often. Thus, this dataset can already provide a reasonable testing scenario for developing more sophisticated hoeing rules. These rules can then either be used on robotic platforms that decide to hoe the ground daily or as part of a farm management system, where the farmer wants to find the best day for hoeing. Additionally, these rules can be extended by a fuzzy system such that conditions that are almost met can be considered as well.

Acknowledgements: The German Research Center for Artificial Intelligence Lower Saxony (DFKI NI) is funded within the framework of Niedersächsisches Vorab by the Lower Saxony Ministry of Science and Culture and the Volkswagen Foundation. The work is funded by the Federal Ministry of Economics and Climate Protection (grant no. 01MK21004A).

Bibliography


[Ch22] Christian Lamprecht: Meteostat Python, meteostat.net. Version 1.6.5. 2022


