Pooled Error Variance and Covariance Estimation of Sparse In Situ Soil Moisture Sensor Measurements

Supplementary Materials

S1. Semivariogram

A semivariogram was constructed to assess the degree of small-scale spatial correlation of SWC within a MZ. The semivariance (γ) for a distance h can be quantified as

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [\theta(x_i) - \theta(x_i + h)]^2 ,$$

where $N(h)$ is the number of data pairs separated by a distance h, while $\theta(x_i)$ is the SWC at location x_i.

Data from field trials with four MZs in 2022 and 2023 at the research center (*Proefstation voor de Groenteteelt*) were used to calculate the semivariances for distances of 1 m, 2 m, and 3 m (Fig. S1). Only measurement times at which all 12 sensors provided data were included, resulting in 65 days of measurements in 2022 and 57 days of measurements in 2023. Within each MZ, the first and second sensors were positioned 1 m apart, the second and third sensors were 2 m apart, and the first and third sensors were separated by 3 m. As a result, each distance (1 m, 2 m, and 3 m) was represented by an equal number of sensor pairs (4). The semivariances were first aggregated by year, combining data from all measurement times and sensor pairs at each specific distance. Subsequently, a weighted average of the semivariances was calculated across the two years, accounting for the varying number of measurement days each year, so that each distance is now represented by 8 sensor pairs.

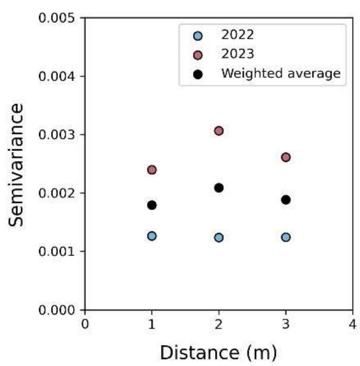

Fig. S1 Semivariogram on MZ scale resulting from two field trials (2022 and 2023). For each growing season, the semivariances were calculated for 1 m, 2 m and 3 m distance, and the weighted average was computed taking into account the measurement days in each year (black).

Interestingly, the small-scale semivariances did not exhibit the expected exponential increase with distance; instead, they remained relatively invariant across the different distances examined. This suggests that SWC is relatively uniform and does not vary significantly with distance at this scale. This uniformity could imply that the factors influencing SWC in the MZs are acting consistently across these small spatial scales. To gain more detailed insights into the spatial variability of SWC within the measurement zones, a denser sensor network, more repeated measurements across a larger area, and additional measurements at distances smaller than 1 m and larger than 3 m would be beneficial.

S2. Irrigation method dependent systematic deviation

The pooled error model was based on a dataset of cropping cycles involving various irrigation methods. The majority of the cropping cycles were irrigated using either a reel irrigation system with an irrigation cannon (46 cropping cycles) or drip irrigation (37 cropping cycles). The remaining cropping cycles used sprinklers (6 cropping cycles) or lateral move irrigation (4 cropping cycles).

Systematic deviations of sensor measurements compared to the 'true' SWC in a MZ measured with a composite soil moisture sample can be expected to vary by irrigation method. A sensor-specific intercept represents these systematic deviations, where a negative intercept corresponds with an overestimation of SWC by the sensor compared to the composite soil moisture sample, while a positive intercept indicates an underestimation. In cropping cycles using drip irrigation, negative intercepts were more common than positive intercepts, suggesting that sensors often overestimated SWC compared to the composite soil moisture sample (Fig. S2). This is likely because the sensors were located close to the dripper lines, within or near the wetting bulb, while soil moisture samples were taken at various locations between the dripper lines.

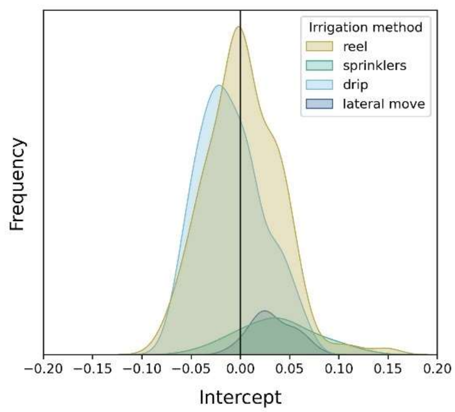

Fig. S2 Density plot of the sensor-specific intercepts grouped by irrigation method. A negative intercept corresponds with an overestimation of SWC by the sensor, while a positive intercept corresponds with an underestimation of SWC by the sensor.

These results indicate that systematic deviations vary depending on the irrigation method used. Therefore, it can be argued that a separate error model for each irrigation method may be necessary. The same might apply for soil type, but this effect could not be assessed in this study as loamy sand (Belgian soil classification: S) heavily predominated the dataset compared to coarser sandy soils as well as finer loamy soils. Developing a separate error model for specific irrigation methods and/or soil types could possibly better address the unique characteristics and deviations of sensors under various conditions, leading to a more accurate understanding of systematic deviations. Nevertheless, this would require an extensive amount of data for each specific condition.