

# The first multi-year publicly available database for Plastic Covered Greenhouse Sentinel-2 Crops (PCGS2Crops)

Manuel A. Aguilar, Abderrahim Nemmaoui, Fernando J. Aguilar, Pablo Cortés, Antonio Jesús Fernández-García

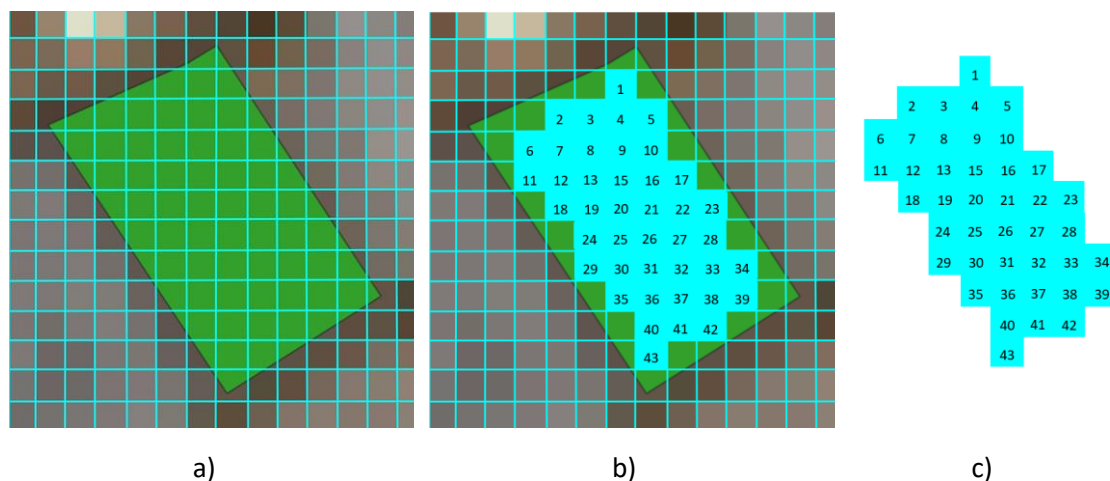
The first multi-year database for Plastic Covered Greenhouse (PCG) Sentinel-2 Crops (PCGS2Crops) is made publicly available to researchers in this repository. PCGS2Crops is based on PCG crop data collected in two agricultural campaigns in Almería (Spain), during the years 2016/2017 and 2024/2025.

The **2016/2017 dataset** was generated through an extensive field campaign with multiple on-site visits to PCGs. We decided to divide the agricultural year into independent campaigns: autumn and spring.

In the autumn campaign, which covered the months of June to December 2016, a total of 1202 PCGs were validated, among which there was a dominant presence of pepper ( $n=511$ ) and tomato ( $n=379$ ), and in a smaller proportion, aubergine ( $n=148$ ) and cucumber ( $n=164$ ). In the spring campaign (from January to June 2017), the database was updated to reflect the transition to cucurbit species, identifying 270 PCGs (out of a total of 1114 PCG) dedicated to melon and watermelon, while the rest of the PCG were grown with pepper ( $n=225$ ), tomato ( $n=394$ ), aubergine ( $n=161$ ), and cucumber ( $n=64$ ).

The **2024/2025 dataset**, mainly based on crop declarations. The autumn campaign (June to December 2024) included a total of 891 PCGs, comprising pepper ( $n=300$ ), tomato ( $n=387$ ), aubergine ( $n=172$ ), and cucumber ( $n=32$ ) (Fig. S3). The spring campaign (January to June 2025) were composed of 1148 PCGs, with 255, 435, 144, 67, and 247 PCGs dedicated to pepper, tomato, aubergine, cucumber, and melon and watermelon, respectively.

**PCGS2Crops** are based on pure pixels, i.e., those completely contained within the polygons of each PCG whose crop we knew (Fig. 1). The idea was to avoid contamination from pixels located at the edges of the PCGs, which might contain spectral signatures of other types.



**Fig. 1.** Pure pixel extraction process using a) the ground truth vector file representing PCGs, b) selecting only the internal pixels totally within the PCG polygons, and c) generating the final dataset.

Each pure pixel of 10 m side within a specific PCG polygon in the PCGS2Crops dataset was indexed by the campaign (i.e., 2016, 2017, 2024 or 2025), the polygon identification code for each PCG, a pixel id for intra-parcel analysis, the PCG crop type label (Pepper=Class 1; Tomato=Class 2; aubergine plus cucumber, called Others=Class 3, and melon and watermelon (M&W)=Class 4. The first three classes were considered in autumn while the four classes were applied in spring), and the reflectance values in percentages corresponding to the 10 bands for each monthly Sentinel-2 L2A composite images. For instance, in 2016 autumn campaign, there were 60 values of reflectance (i.e., 10 bands x 6 monthly composite images) and only three labels for crops because there are no melon and watermelon at this season. Thus, the reflectance values depicted in the PCGS2Crops for June, 2016 were named as Blue\_mean\_6 (B2), Green\_mean\_6 (B3), Red\_mean\_6 (B4), RE1\_mean\_6 (B5), RE2\_mean\_6 (B6), RE3\_mean\_6 (B7), NIR\_mean\_6 (B8), N\_NIR\_mean\_6 (B8A), SWIR1\_mean\_6 (B11), and SWIR2\_mean\_6 (B12). For more information, consult the work entitled: **Using artificial intelligence and Sentinel-2 time series for greenhouse horticultural crops classification**.

Table 1 depicts the total number of PCGs whose crops were monitored during the four campaigns of our study, and which were stored in vector files. In addition, Table 1 shows in brackets the number of pure pixels representing a given crop that were extracted from the monthly Sentinel-2 images. It is important to note that, in the four campaigns, the number of PCGs containing aubergine or cucumber was much lower than the number dedicated to peppers, tomatoes, or cucurbits (M&W). Because of this, it was ultimately decided to group aubergine and cucumber crops into a category called Others.

**Table 1.** Number of PCGs and pure pixels (in brackets) controlled in each of the four campaigns in relation to the crop growing inside. Others: Aubergine and Cucumber. M&W: Melon and watermelon.

| Crop   | Autumn 2016  | Spring 2017  | Autumn 2024 | Spring 2025  |
|--------|--------------|--------------|-------------|--------------|
| Pepper | 511 (26446)  | 225 (11639)  | 300 (13682) | 255 (12109)  |
| Tomato | 379 (20441)  | 394 (20999)  | 387 (18451) | 435 (21053)  |
| Others | 312 (15567)  | 225 (11186)  | 204 (8946)  | 211 (9670)   |
| M&W    | -            | 270 (15403)  | -           | 247 (12684)  |
| Total  | 1202 (62454) | 1114 (59227) | 891 (41079) | 1148 (55516) |

The ground truths for these four campaigns were curated and presented in four Comma-Separated Values (CSV) called:

2016\_Autumn.csv

2017\_Spring.csv

2024\_Autumn.csv

2025\_Spring.csv