Article

PPMaP: reproducible and extensible open-source software for plant phenological phase duration prediction and mapping

Henri E. Z. Tonnang [1,2,3,*], Ritter A. Guimapi[2], Bruce Anani[1], Dan Makumbi[1], Bester T. Mudereri[3], Tesfaye Balemi[4] and Peter Craufurd[1]

[1]International Maize and Wheat Improvement Center (CIMMYT), off UN Avenue, Gigiri, ICRAF House, P.O. Box 1041-0062, Nairobi, Kenya; a.bruce@cgiar.org (A.B.); d.makumbi@cgiar.org (D.M.) p.craufurd@cgiar.org (P.C.)

[2]NIBIO-Norwegian Institute of Bioeconomy Research, P.O. Box 1433 Ås, Norway, ritteryvan@gmail.com (R.A.G.)

[3]International Centre of Insect Physiology and Ecology (*icipe*), P.O. Box 30772-00100, Nairobi, Kenya; bmudereri@icipe.org (BTM)

[4]Ethiopian Institute of Agricultural Research, P.O.Box 2003, Addis Ababa, Ethiopia; tesfayeb2005@yahoo.co.uk (T.B)

*Correspondence: htonnang@icipe.org (H.E.Z.T.)

Abstract: Understanding the detailed timing of crop phenology and their variability enhances grain yield and quality by providing precise scheduling of irrigation, fertilization, and crop protection mechanisms. Advances in information and communication technology (ICT) provide a unique opportunity to develop agriculture-related tools that enhance wall-to-wall upscaling of data outputs from point-location data to wide-area spatial scales. Because of the heterogeneity of the worldwide agro-ecological zones where crops are cultivated, it is unproductive to perform plant phenology research without providing means to upscale results to landscape-level while safeguarding field-scale relevance. This paper presents an advanced, reproducible, and open-source software for plant phenology prediction and mapping (PPMaP) that inputs data obtained from multi-location field experiments to derive models for any crop variety. This information can then be applied consecutively at a localized grid within a spatial framework to produce plant phenology predictions at the landscape level. This software supports the development of process-oriented and temperature-driven plant phenology models by intuitively and interactively leading the user through a step-by-step progression to the production of spatial maps for any region of interest. Maize (*Zea mays* L.*)* was used to demonstrate the robustness, versatility, and high computing efficiency of the resulting modeling outputs of the PPMaP. The framework is implemented in R, providing a flexible and easy-to-use GUI interface. Since this allows appropriate scaling to the larger spatial domain, the software can effectively be used to determine the spatially explicit length of growing period (LGP) of any variety.

Keywords: Plant development rate, temperature-dependent, landscape, multi-location trials

1. Introduction

The recent rise and evolution of data analytics have greatly modernized the worldwide agriculture environment thereby improving the efficiency of monitoring farming environments [1,2].

Mathematical and statistical modeling tools embedded within computer programs and packages have been used intensively over the last decade with the intent to enhance agricultural production at scale [3–8]. The output of such technologies has revolutionized the empirical optimization of production and accurate predictions of output which has aided in precision agriculture planning and management [7,9,10].

This paper contributes a reproducible and practical software capable of providing plant phenological phase period prediction and mapping (PPMaP). The PPMaP package, freely available at https://github.com/Atoundem/PPMaP, combines data obtained from multi-location field experiments for plant phenology to derive temperature-dependent models for any crop variety. This information is then applied consecutively at an individual grid-level within a spatial framework to produce predictions at scale [8]. Unlike the majority of the available plant analysis tools such as the North Carolina State University/Animal and Plant Health Inspection Service Plant Pest Forecasting System (NAPPFAST) [9] that uses the degree-day, this software supports the development of process-oriented temperature-driven plant phenology models and interactively leads the user through a step-by-step progression until the production of a spatial ASCII (American Standard Code for Information Interchange) file for any region of interest.

The modules with which the PPMaP software was developed were cognizant of the nexus between stages of plant development and the various variables that influence each stage. Plant development denotes the scheduling of occurrences in the life cycle of a plant that can be described as the increase in weight, volume, length, or area of some of the parts or the whole plant [11]. These plant growth and developmental processes are used to delineate the phenological stages of a crop. However, for modeling purposes, it is essential to separate the two main processes i.e. rate of biomass accumulation and the length of growth, as they are affected by different environmental variables even within intra-species [12,13]. On one hand, the rate of biomass accumulation is primarily determined by the amount of light captured by plants over a range of temperatures, whereas the length of growth for a specific variety is typically dependent on daily temperatures. It has been insinuated that the probable rate of biomass growth is comparatively constant over space and time when temperature values are within the range for plant growth, while the period of growth, changes in space and time [11,12].

Naturally, plants are considered as chemical "machines" that are sensitive to the temperature that characterizes the environment in which they are grown. In this context, earlier research has reported that each mechanism of plant development (enzyme reaction, metabolic sequence, and physiological process) is temperature-dependent [14–17]. Therefore, it follows that temperature determines the progression of processes to the next developmental stage and their respective rate of development [18]. Numerous relationships have been established to portray the way temperature influences plant development [19–21]. These comprise degree-days, day-degrees, heat units, heat sums, thermal units, and growing degree-days [8,16,17]. Depending on the geographic location, there is a dissimilar weighting to the night and day periods for a specific crop, hence a variety might have different periods of growth as determined by these locational temperature variations. Additionally, this is also in consideration of the specific developmental characteristics of a crop, for instance, maize (*Zea mays* L.) which stops development when it reaches physiological maturity regardless of the success in development [22]. Therefore, the formulation of mathematical expressions to represent the development of plants and insects has evolved with emphasis placed on determining the lowest temperature at which development is zero (base temperature (Tb)) and the optimum temperature (To) above which development ceases to increase or begins to decline.

Motivated by this postulate, several models have been formulated towards enhancing understanding of these phenomena, for which each model has its strengths and weaknesses [13,19,20]. In general, relationships have been established between the time of development (duration between life-phases) of plant and temperature either empirically or through process-based methods. Process-based equations are formulated and parameterized with the biological knowledge of the plant, while empirical models only capture the trends of the data [20]. Despite the abundance of literature on the use of temperature-dependent models for the prediction of plant developmental

phases [20,22,23], no simple tool exists to help develop and project the phenological phases of plants at the landscape level. Besides, the majority of web services applications and analytics, which are currently available, are often tailored for the developed world with nothing to support the majority of farmers in developing countries [9].

For instance, the complex role of genetic factors in determining the performance of genotypes in different environments and selecting superior genotypes in target environments was demonstrated by Dia et al. (2016) [7]. In their study, the Genotype x Environment interaction (G×E) analysis was applied to describe the stability of genotypes and the response of crop in a multiplication context. This method helps to estimate the statistical parameters which guide on the selection of genes that have more adapted environmental and agronomic traits [7]. Additionally, other several applications and online tools for plant image identification, analysis, and sharing of leaf images to interpret and predict potential yield have been proposed [10,24,25]. All these tools provide mechanisms to handle large datasets and further facilitate the distribution of information among the growing scientific community. Additional examples of efforts and software that were developed to improve agricultural modeling include the CN-CLASS [21] that was initially developed to study Carbon (C) stock in forest ecosystems but was later modified to link crop phenological development and C allocation during the growth of maize. Also, CropScape is a web service-based application for discovering and disseminating geospatial cropland data products for decision making [4]. Other software include the LEAF-E that was developed to analyze plant leaf growth through function fitting, while TIPS is an automated computer system for processing image-based phenotyping of maize tassels [26]. However, the weakness of all these methods is that they generalize and facilitate the extrapolation of the potential areas for the optimum growth of the variety based on the known genotype, thus missing the fundamental location-specific parameters.

Therefore, we introduce new software that solves the limitations of the existing phenology determining platforms. The PPMaP is a framework that enables multi-location field experiments for plant phenology to derive temperature-dependent models for any crop variety and then provide a grid specific recommendation domain for the variety. This enables the successful extrapolation of the length of growing periods (LGPs) for any varieties with high levels of accuracy with the lowest potential risk of failure [14]. This software also counters the traditional methods in which breeders find the best varieties by planting them in a diverse set of locations to measure performance over many seasons. The number of locations that these breeders can test the new varieties is however limited and can cause uncertainty when trying to choose the best genotypes for farmers [27,28]. Thus using PPMaP breeders can accurately predict the performance of each genotype in untested environments and therefore make better decisions on which genotypes to move forward and provide to farmers [14]. Ultimately this results in better decision making and increased crop production to help address food security threats at a global level. In this current study, the phenology model was developed for maize to demonstrate resulting modeling outputs using data collected in high maize producing regions in Ethiopia and Nigeria. In sub-Saharan African, maize stands out as an essential staple food, providing approximately 25% of total calories in the average diet of the majority of the population [29,30].

2. Materials and Methods

The software analytics used in the development of PPMaP was founded on the following four principles: i) local, regional and global relevance, ii) representativeness of the major agro-ecological zones within the region of interest, iii) dependence on high attribute and smallest datasets, and iv) adequate validation of the results with field trial results.
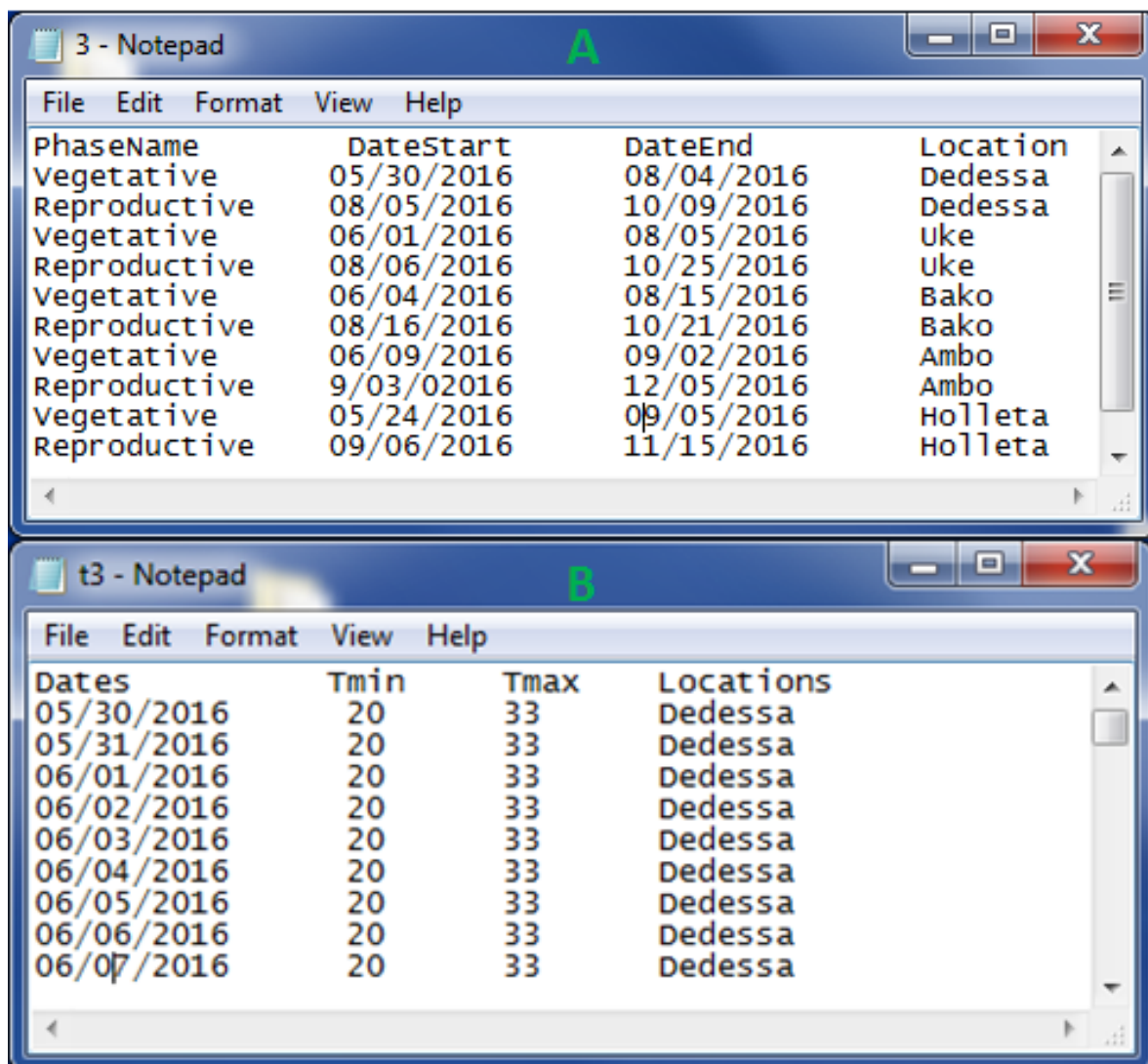
2.1 PPMaP input datasets

2.1.1 Plant phenology datasets

The cycle of a plant is split into different development stages or phases. Maize (crop used as to test the software) growth is divided into the vegetative (VP) and the reproductive (RP) phases. The

VP commences at emergence, goes through the *nth* leaf, and ends with tasseling and flower [15]. The RP starts at flowering and is comprised of silking, blister, milk, dough, and dent stages, and stops at physiological maturity, which is reached when all kernels have maximum dry weight [31].

Input data into PPMaP for modeling the developmental rate of a plant are the duration of the individual phase of phenology. The data are collected from multi-locations characterized by a large temperature gradient to estimate how the duration of each phenological phase responds to diverse growing conditions. Main observations to be recorded include the sowing date, days to emergence, days to flowering, and days to physiological maturity. Days to emergence can be recorded by counting the number of maize plants daily while the emergence date is when 50% of the plants in the plot have emerged above the soil surface. A few days after the emergence of all plants, many plants can be randomly selected and tagged for use to record the duration of the VP and RP.   The flowering date is recorded when 50% of tagged plants produced male or female flowers. Similarly, days to physiological maturity are recorded when 50% of tagged plants attain physiological maturity. A sample of how the data are organized before inputting into the software is shown in Figure 1.

Figure 1.(A) Example of developmental phase start and end dates of a maize variety in five locations (Dedessa, Uke, Bako, Ambo, and Holleta) of Ethiopia, and (B) captures the average values of maximum and minimum temperatures in the five locations. Both files are saved as a txt-file (tab-delimited) and are used as input to PPMaP software. In (A), the first column indicates the developmental phase (vegetative or reproductive) of maize, the second column indicates the starting date of the phase, and the third column is the end date of the phase. Column four is the location in which the experimental trial was conducted. In (B), column one is the date of capture, column two,

and three the minimum and maximum values of temperature respectively, while column four is the name of the location where the trails were conducted.

2.1.2 Temperature datasets

The software uses two distinct sets of temperature. The first set of temperatures is directly obtained from the weather stations within the vicinity of each field experiment. Minimum and maximum temperature for the trial period are required. The information is processed to yield the average seasonal temperature for the site during crop growth. It is associated with the duration of each plant phase when loading into the software (Figure. 1). The second type of dataset needed by PPMaP for landscape mapping is the minimum and maximum weekly/monthly temperature organized in raster files with a *.flt* file format. The data can be obtained from any climate-based repositories and databases. To demonstrate the functionalities of the tool applicable to tropical maize, temperature data were extracted from the Worldclim website (www.worldclim.org). These datasets are presented in raster with "Float" file format (*.flt* and *.hdr* file) and consist of sets of climate layers (grids) with a spatial resolution ranging from 30 arc-seconds (~ 1km) to 10 arc-minutes.

2.2 *PPMaP software components and data processing mechanisms*

2.2.1 PPMaP modeling component

The modeling component is comprised of 82 factor-process based development rate functions obtained from literature; which have been applied in agricultural production, either in the context of insect or crop phenology modeling [13,20,32]. A plant developmental rate model predicts the average proportion of development at a certain temperature or the fraction of development completed per unit time. These fractions are then accumulated under field conditions and are treated as independent variables during the factor-based modeling formulation. Model expressions in the PPMaP database are classified by the form of relationship and structure (linear, logistic, exponential, sigmoid, logarithmic, polynomial, and square root expressions).

Modeling with PPMaP starts with data standardization by dividing the development period of the individual phase by the mean development period of each field experimental site average temperature. This allows data obtained from multiple location trials with similar average temperature values to be merged for unique scrutiny. Because the direct selection of an appropriate mathematical expression to represent the development phase of a crop is challenging [13,20,33], PPMaP software model component uses its in-built statistical criteria to compare, select and display equations with the respective parameters and graph. Parameter values are estimated by fitting model equations to the field data. The Levenberg-Marquardt (LM) algorithm [34] is implemented in the software for the estimation of the model parameters. This algorithm combines both the 'steepest descent' and the Gauss-Newton method to appraise equations. It works iteratively to find the minimum of a function expressed as the sum of squares between the nonlinear model output and the observed datasets [34]. The PPMaP software offers an interactive process, in which initial parameter values for the selected model are altered; followed by the matching of model output to observed data. After the initialization of the models, the LM algorithm is launched and the goodness of fit procedure is used to find the parameters of the model. The model parameter comparison is done by the coefficient of determination adjusted r-squared (R^2_Adj) [13,20]; while the Akaike's Information Criterion (AIC) and the Model Selection Criterion (MSC) [20] are used to choose the model that best fits the input datasets (Figure. 2).
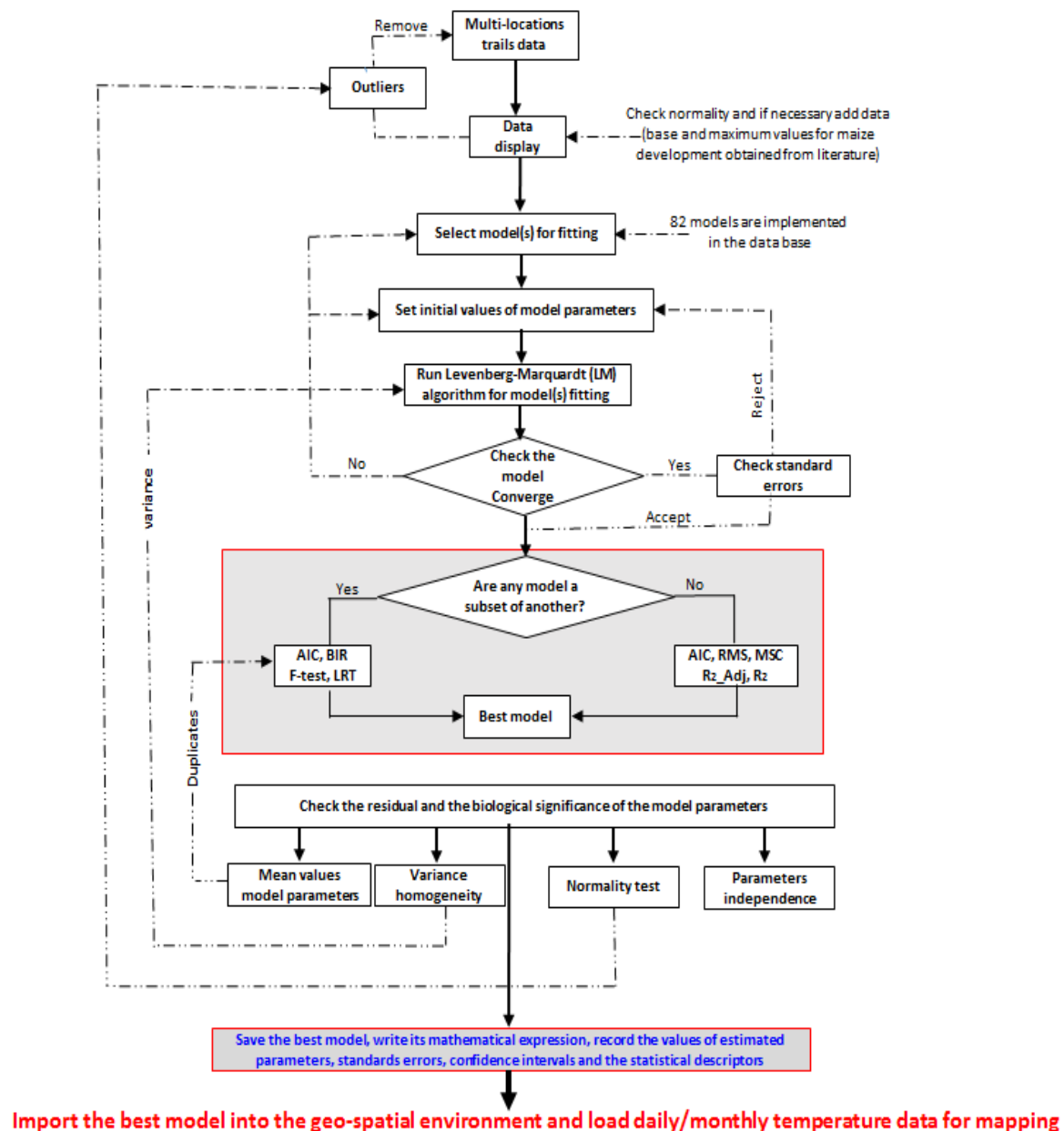
Figure 2: Flow chart demonstrating the steps within the PPMaP software from fitting plant phenology data obtained from multi-trails location to model selection and mapping. The thick, thin and dashed arrows are the essential steps, substeps, and feedback respectively. Akaike's Information Criterion (AIC), Model Selection Criterion (MSC), Likelihood Ratio Test (LRT), Bayesian Information Criterion (BIC) (Adapted from Archontoulis and Miguez, 2015).

2.2.2 PPMaP mapping component

A shapefile representing the boundaries of the region of interest is divided into grids and loaded into the software. The model obtained during the modeling step is run at each grid of the raster file. The gridded temperature data is also loaded and with an in-built function, the values are extracted from the database simultaneously and then arranged in a matrix format using longitude as a column and latitude as a row. A point object picks the model representing the plant development rate for a specific phase, and then sequentially replaces the variable temperature in each grid by the value obtained from the database. After computation in each grid, a new matrix is generated with the values of the development rate in the respective grid of the raster file. These values are then inversed (1/development rate) to estimate the number of days used to complete each phase of plant

development. The results are converted into ASCII files and can be transferred to any geographical information system (GIS) [35] software for better visualization.

2.3 PPMaP software implementation

The following two programming methods were encompassed and combined to implement the PPMaP software: (i) a procedural approach that employed the R [36] constituent for model fitting and estimation of statistical criteria [37], and (ii) Java object-oriented programming language was used to interconnect the software components and develop the graphic user interface (GUI). This was developed on the Eclipse platform [38]. The R constituent of the PPMaP comprises of R-scripts used to code mathematical expressions (models), processes for parameter estimation and model selection, and the mapping environment. Many R packages such as 'minpack.lm' [39], 'MASS' [40]; 'maptools' [41], and 'maps' [42] are employed to handle nonlinear regression analysis and model fitting and to carry out spatial operations leading to the generation of maps.

The Java object-oriented programming allows the development of a rapid and reliable end-user interface with objects, subclasses of image, color, and font integrated into the Eclipse workbench. The 'Rserve' is a tool that is used to evaluate R codes and ensure their integration into the Java application [43]. The 'Rserve' was applied within the PPMaP software to ensure two-way communication between the R constituent and the Java object-oriented codes. When a request is made, 'Rserve' unlocks the link to collect details of the request, and then creates another link with the R constituent to transfer information for execution and sends back the result to the GUI for display. The GUI offers a direct connection between the user and PPMaP software system. Numerous utilities are developed to bring PPMaP functionalities together in a stop screen layout to make it operate intuitively. The PPMaP software iteratively helps the user to create a project (Figure. 3); import, modify, plot, and visualize data (Figure. 4); carry out modeling (Figures. 5 and 6), and to project the temperature-dependent model at scale (Figure. 7).
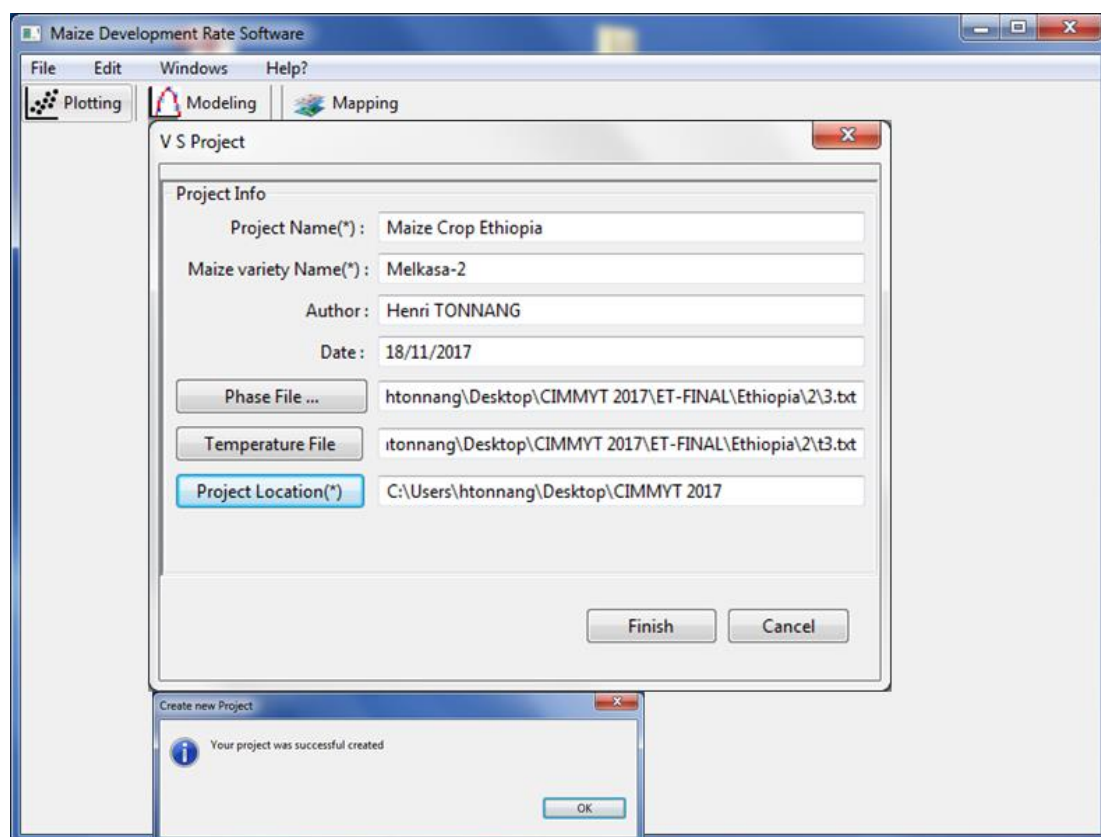
Figure 3: Display of the PPMaP software window for project creation. Under project info, there is the project name (Maize Crop Ethiopia), the maize variety name (Melkasa-2), the author (Henri TONNANG), and the date (18/11/2017) in which the project was created. Phase file designates the path to load file **A** (Figure. 2). The temperature file is the path to load file **B** (Figure. 2). Project location designates the folder in which all project files are stored. A well-created project prompts the software to display the message "*Your project was successfully created*".
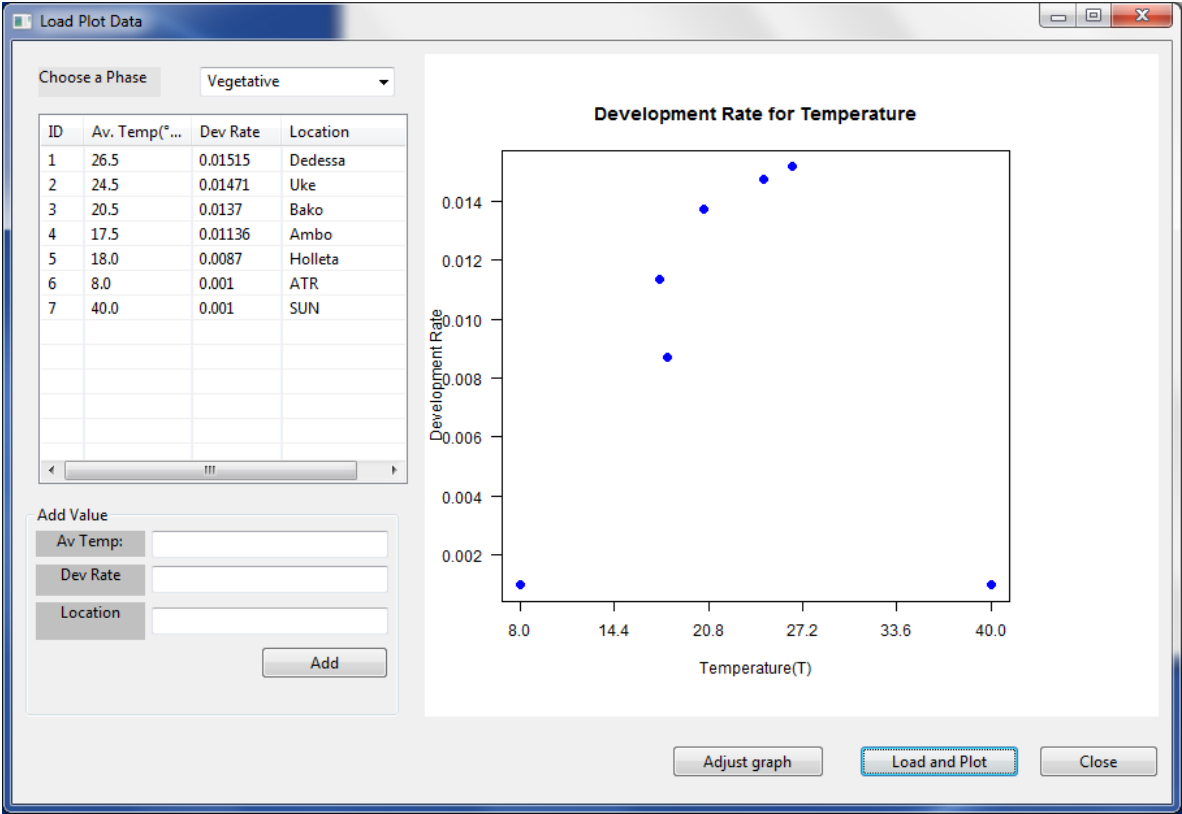

Figure 4. PPMaP data file display window. It contains the compilation of data of the vegetative phase of a maize variety. The columns on the left box contain, the ID, average temperature, numerical values of the development rate, and the locations. The right box shows in two dimensions the data recorded in the left box.
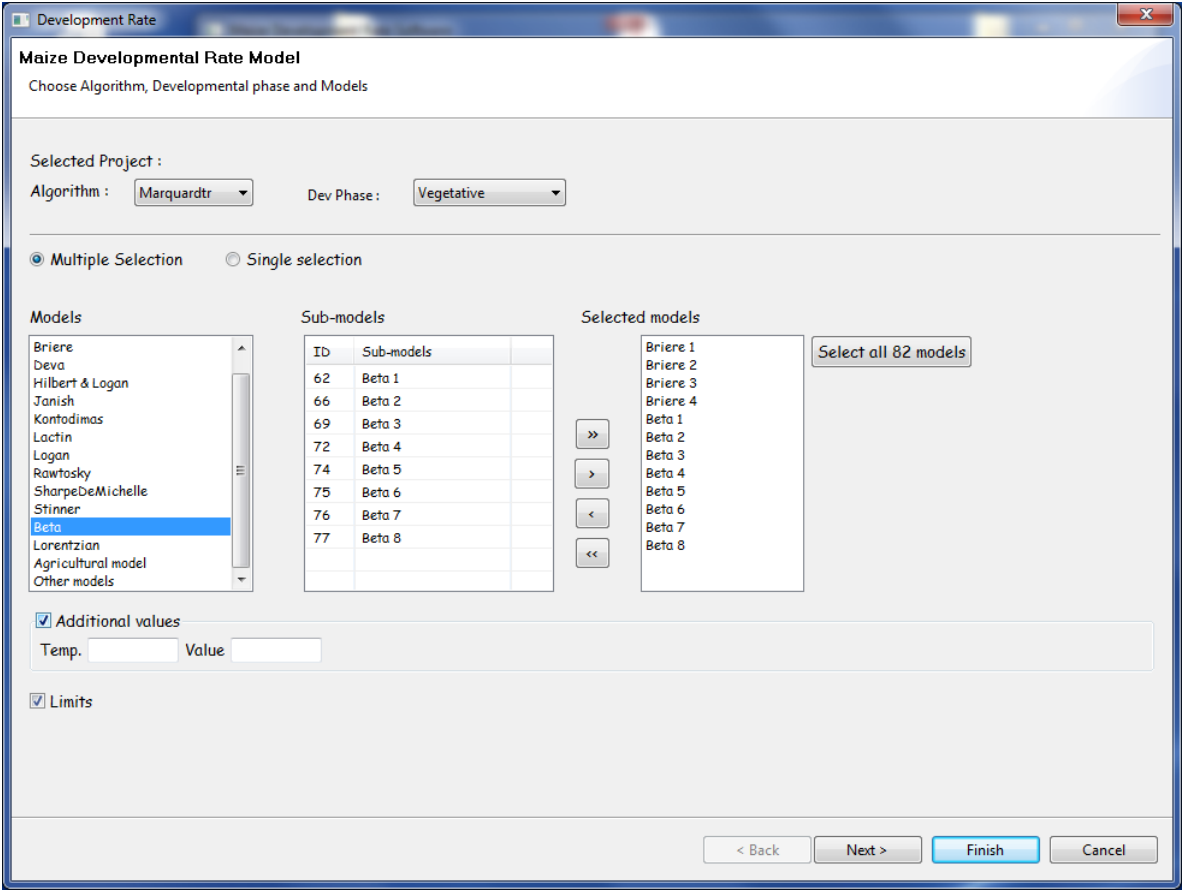
Figure 5. Illustrates the user interface of the "modeling component" of the PPMaP software. It shows the library of mathematical expressions that are used to describe the relationship between plant phase development and temperature. The left box contains the models characterized by the name of the author that pioneers the formulation. The center box contains the sub-models, which are derivative of the models from the left box. The right box contains the models selected for fitting. For example, in Briere$_i$, the index i represents the number of the derived model from the original Briere model.
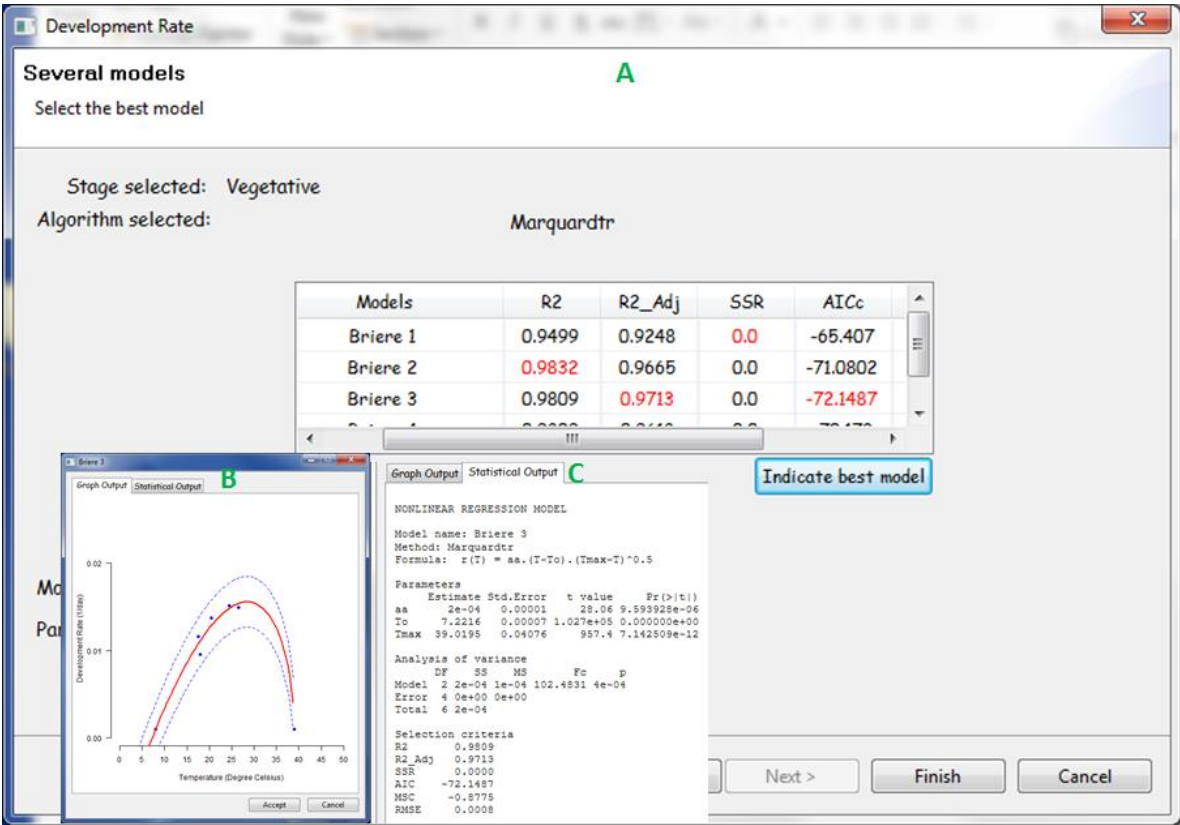
Figure 6. (A) Window of the PPMaP software presenting diverse models, each with corresponding model selection criteria for comparison; usually the best model has the smallest value of Akaike's Information Criterion (AIC). (B) Example of points of data (blue) for a variety of maize at the vegetative phase fitted by Briere temperature-dependent mathematical expression (red curve). (C) The statistical criteria for model selection, the model parameters, and the (ANOVA) analysis of variances are shown.
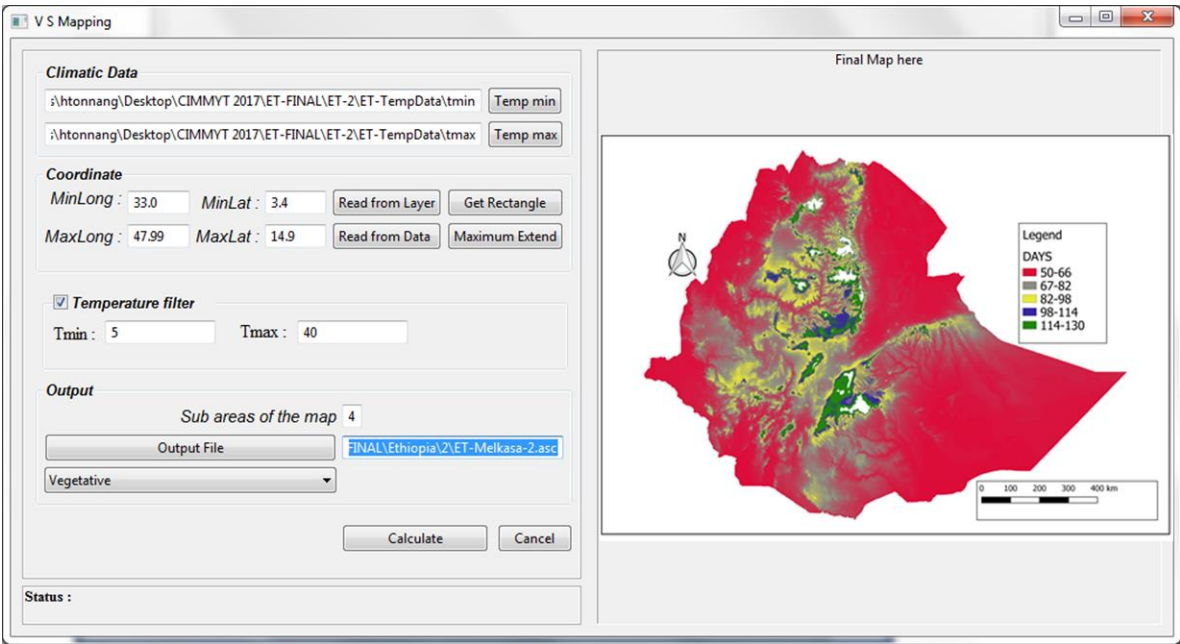

Figure 7. PPMaP window showing the mapping section of the software. The paths for inputting minimum and maximum temperatures are shown. Below these paths are the geographical coordinates of the region chosen for mapping. The map of Ethiopia showing the number of days a

maize variety could spend during the vegetative phase appeared in a box at right. Depending on the location, this maize variety can spend from 50 to 130 days to start flowering.

2.4 PPMaP model validation and performance at scale

PPMaP software considers the validation as the level to which a selected model rightly predicts the duration of an individual phase of the plant when compared to experimental data. To carry out this procedure, independent data acquired from various field trial sites that were not used during model development and calibrations are recommended for use. The temperature at point location is replaced in the expression of the inverse of development rate to estimate the duration of the individual phase of the plant, and the result is compared with actual recorded data from the field trial. A well-validated model is used to measure the performance at scale.

Figure. 8 shows the phenology maps of two maize varieties (MV-1 and MV-2) that are grown both in Ethiopia and Nigeria respectively. The period to complete the individual phase during the growing period in each country differs considerably due to variability in the temperatures. The growing period of variety MV-1 in Nigeria is predicted between 40 to 50 days, whereas in Ethiopia it could take from 50 to 130 days. MV-2 also presents a similar difference in the growing period in Nigeria (59 – 64 days) and Ethiopia (60 to 170 days). These results demonstrate the high level of variability that exists in agro-ecological areas in an individual country. It is noted that both maize varieties (MV-1 and MV-2) can successfully grow in most parts of Nigeria and Ethiopia. However, the time each variety will take to reach maturity will differ from one location to another. With Ethiopia having wider variability in altitude compared to Nigeria, this translates to diverse temperature regimes and the time taken by both maize varieties to reach maturity is longer than in Nigeria. At higher altitudes, the temperatures are lower, making the duration of development longer than at lower altitudes.
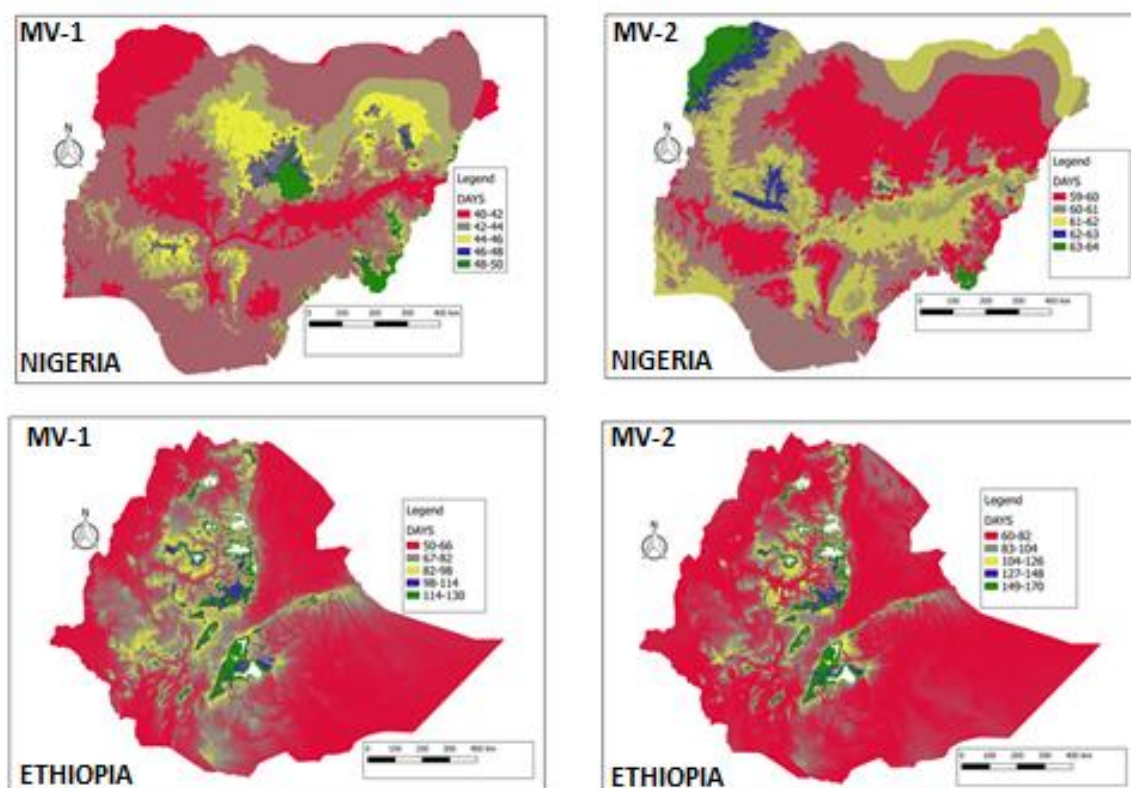

Figure 8. Maps of Nigeria and Ethiopia showing the number of days 2 maize varieties (MV-1 and MV-2) would take during each phase (vegetative or reproductive) of development in different agro-ecologies. The small white dots are areas with missing values of temperature

3. PPMaP software usefulness and discussion

Phenology is an area of research that assesses the cyclical recurrence of events in the growth and development of plants and insects [12]. Thus, predicting plant phenology is important as it can be used to improve crop productivity. Precisely, forecasting stages of crop development duration is needed to find genotypes with the desired growth period that allows growers to optimize yields. Adequate forecast of plant phenological events further guides in management decision making like the timing of pesticide application, scheduling the timely harvest of crops, or synchronizing the flowering of cross-pollination crops for seed production [11].

Notwithstanding the multitude of literature [11,12,20] on the use of temperature-dependent models for modeling plant developmental phases, no software can help non-modelers to establish mathematical expressions for useful prediction of plant phenological events. The PPMaP software presented here is a unique computer-aided tool capable of analyzing multi-location field trial information of plants to yield complete phenology models. PPMaP encompasses a full library of mathematical expressions and statistical criteria that guide users to discovery and to select the best model to describe the temperature dependence of the development for each phase of the plant. The starting point in developing plant phenology models with PPMaP is data standardization and grouping according to stage and site-specific characterized by the average temperature in which the trials were conducted. Fitting the developmental rate curves directly to the reciprocal of developmental duration follows this step. To account for variability among plants, PPMaP applies the distributional concept demonstrated in [44,45], which assumes that variation in developmental rates at constant temperature is the direct result of variability in the concentration of rate-controlling enzymes while the developmental rate variance is proportional to the mean developmental rate.

An important attribute of PPMaP is the relative ease and speed with which the overall process of producing the model and projecting at scale to yield the phenology map of the crop is done. The software environment makes programming and coding unnecessary, and hence allows the estimates of the duration of plant phases effortlessly available to researchers. PPMaP has a graphical user interface and a collection of appropriate models for representing different processes that determine the plant life cycle. Selection is based on analytic and adequate evaluation criteria. Fitting mathematical expressions using either maximum likelihood estimation (MLE) or least-squares estimation (LSE) and treatment medians is difficult and sometimes often leads to over parameterization and convergence [20]. Therefore, PPMaP offers a solution to the problem whereby a user may be requested to increase the number of data points or use the in-built software features that allow the use of censored data to facilitate the analysis. In some cases, model convergence with certain initial parameters may remain an issue. However, PPMaP provides an option to automatically modify the initial parameters of models to ease the convergence of the fitting algorithm. However, fitting mathematical expressions remains a complicated task and should be done carefully. There is no standard method to choose among competing equations; that is, the decision on choosing a model should not be left to a computer program, and thus the user of PPMaP needs to critically consider and identify an appropriate equation that best delivers biologically meaningful and statistically significant parameters suited to their objective and species or variety [20].

In Rykiel, (1996), validation is defined as the degree to which a model output matches independent data sets. Thus, PPMaP uses the same concept and has a mechanism for using independent datasets (not used for model development and calibration) obtained from diverse field trials and sites to make predictions and compare results. In this context, validation is an important instrument for scaling life cycle events of plants over a wide spatial and temporal range and for examining the crop reactions to changing or novel climatic conditions, and it provides some level of understanding and adaptation in different agro-ecological zones.

Taking crop research to scale requires a change in the way studies are conducted as well as the extent to which essential advice to farmers can be provided. Two types of frameworks exist the top-down approach consisting of analyzing grid cells to cover a wide area and then downscaling the results, and the bottom-up approach in which several sites representing major crop production areas are selected for analysis followed by a spatial upscale [47]. Consequently, PPMaP software supports

the concept of researching to the appropriate scale by using the bottom-up approach. Inputting phenological data of plants to carry out analysis with geospatial representation is thus a critical step to tailoring solutions at the landscape level. Considering the diversity of agro-ecological zones where crops are cultivated, we argue that it is unproductive to conduct research studies on plant phenology without providing means of upscaling results to landscape-level while still safeguarding field relevance. Herein, we presented software embedded with a method that allows appropriate scaling to larger spatial domain research findings for decision-making. Global warming might be affecting the phenology of plants in their respective natural environments.

Tools based on hyperspectral imaging exist for differentiating crop phenology and seed varieties [10,24]. Plant phenotyping [10,27] is an emerging field of research that links genomics with plant ecophysiology and agronomy. It demonstrates that the functional plant body (phenotype) is produced during plant growth and development from the dynamic interaction between the genetic setting (genotype) and the physical area in which plants develop (environment). The understanding of the connections and feedbacks mechanisms existing between phenotype, genotype, and the environment is paramount for increasing crop productivity, therefore, we hope that PPMaP which helps to understand how genotype can respond to environmental conditions will contribute towards improving knowledge on this topic.

In several countries of sub-Saharan Africa (SSA), information on crop seed packages is often incomplete; crop varieties are classified by altitude (low, medium, and high) and maturity groups (early, intermediate, and late). PPMaP is a tool that could help provide location specific and near accurate length of crop growth duration of different varieties. Consequently, we propose to compact the phenology maps of different maize varieties produced by PPMaP, add features such as the name, type, color, potential yield, maturity class, ecology, resistance, and tolerance level to pests/diseases for individual varieties to develop applications. This software can be used to enhance the capacity of extension and agribusiness service providers (agro-dealers) in providing recommendations at various scales especially in countries where the landscape varies considerably in altitude such as Ethiopia. When a new variety is developed, the norm is to conduct multi-environmental trials to evaluate the performance of the variety before release. Such trials are very expensive and PPMaP can be a starting point for reducing the number of experimental sites and thus associated costs.

4. Conclusions

This study has demonstrated that PPMaP is a tool that has vast potential to provide a location-specific and near accurate length of crop growth duration of different varieties by supporting the concept of researching to the appropriate scale. It is an object-oriented reproducible and extensible framework for phenology mapping at scale. Besides, PPMaP provides an easy-to-use comprehensive framework to perform the entire modeling process without the extensive nee for programing skills hence user friendly to potential users of any computer skills level. The software is thus designed to enable users to extend it and share the new data, methods, or procedures to reproduce them by other users.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The appendix is an optional section that can contain details and data supplemental to the main text. For example, explanations of experimental details that would disrupt the flow of the main text, but nonetheless remain crucial to understanding and reproducing the research shown; figures of replicates for experiments of which representative data is shown in the main text can be added here if brief, or as Supplementary data. Mathematical proofs of results not central to the paper can be added as an appendix.

Appendix B

All appendix sections must be cited in the main text. In the appendixes, Figures, Tables, etc. should be labeled starting with 'A', e.g., Figure A1, Figure A2, etc.

References

1.	Saiz-Rubio, V.; Rovira-Más, F. From smart farming towards agriculture 5.0: A review on crop data management. *Agronomy* **2020**, *10*.

2.	Mendes, J.; Pinho, T.M.; Neves dos Santos, F.; Sousa, J.J.; Peres, E.; Boaventura-Cunha, J.; Cunha, M.; Morais, R. Smartphone Applications Targeting Precision Agriculture Practices—A Systematic Review. *Agronomy* **2020**, *10*, 855.

3.	Gungula, D.T.; Kling, J.G.; Togun, A.O. CERES-Maize predictions of maize phenology under nitrogen-stressed conditions in Nigeria. *Agron. J.* **2003**, *95*, 892–899.

4.	Han, W.; Yang, Z.; Di, L.; Mueller, R. CropScape: A Web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support. *Comput. Electron. Agric.* **2012**, *84*, 111–123.

5.	Joly, A.; Goëau, H.; Bonnet, P.; Bakić, V.; Barbe, J.; Selmi, S.; Yahiaoui, I.; Carré, J.; Mouysset, E.; Molino, J.F.; et al. Interactive plant identification based on social image data. *Ecol. Inform.* **2014**, *23*, 22–34.

6.	Bobryk, C.W.; Myers, D.B.; Kitchen, N.R.; Shanahan, J.F.; Sudduth, K.A.; Drummond, S.T.; Gunzenhauser, B.; Gomez Raboteaux, N.N. Validating a digital soil map with corn yield data for precision agriculture decision support. *Agron. J.* **2016**, *108*, 957–965.

7.	Dia, M.; Wehner, T.C.; Arellano, C. Analysis of genotype × environment interaction (G×E) using SAS programming. *Agron. J.* **2016**, *108*, 1838–1852.

8.	Tonnang, H.E.Z.; Balemi, T.; Masuki, K.F.; Mohammed, I.; Adewopo, J.; Adnan, A.A.; Mudereri, B.T.; Vanlauwe, B.; Craufurd, P. Rapid acquisition, management, and analysis of spatial Maize (Zea mays L.) phenological data — Towards 'Big Data' for agronomy transformation in Africa. *Agronomy* **2020**, *10*.

9.   Magarey, R.D.; Fowler, G.A.; Borchert, D.M.; Sutton, T.B.; Colunga-Garcia, M.; Simpson, J.A. NAPPFAST: An internet system for the weather-based mapping of plant pathogens. *Plant Dis.* **2007**, *91*, 336–345.

10. Inostroza, L.; Acuña, H.; Munoz, P.; Vásquez, C.; Ibáñez, J.; Tapia, G.; Pino, M.T.; Aguilera, H. Using aerial images and canopy spectral reflectance for high-throughput phenotyping of white clover. *Crop Sci.* **2016**, *56*, 2629–2637.

11. Ritchie, J.T.; Nesmith, D.S. Temperature and Crop Development. In *Modeling Plant and Soil Systems-Agronomy Monograph no. 31*; Madison, USA, 1991; pp. 5–29.

12. Schröder, W.; Schmidt, G.; Schönrock, S. Modelling and mapping of plant phenological stages as bio-meteorological indicators for climate change. *Environ. Sci. Eur.* **2014**, *26*, 1–13.

13. Wu, L.; Feng, L.; Zhang, Y.; Gao, J.; Wang, J. Comparison of five wheat models simulating phenology under different sowing dates and varieties. *Agron. J.* **2017**, *109*, 1280–1293.

14. Tonnang, H.E.Z.; Makumbi, D.; Craufurd, P. Methodological approach for predicting and mapping the phenological adaptation of tropical maize (Zea mays L.) using multi-environment trials. *Plant Methods* **2018**, *14*, 1–12.

15. Liu, Y.; Xie, R.; Hou, P.; Li, S.; Zhang, H.; Ming, B.; Long, H.; Liang, S. Phenological responses of maize to changes in environment when grown at different latitudes in China. *F. Crop. Res.* **2013**, *144*, 192–199.

16. Sacks, W.J.; Kucharik, C.J. Crop management and phenology trends in the U.S. Corn Belt: Impacts on yields, evapotranspiration and energy balance. *Agric. For. Meteorol.* **2011**, *151*, 882–894.

17. Lobell, D.B.; Bänziger, M.; Magorokosho, C.; Vivek, B. Nonlinear heat effects on African maize as evidenced by historical yield trials. *Nat. Clim. Chang.* **2011**, *1*, 42–45.

18. Hopkins, W.G. *Plant development*; Infobase Publishing.: New York, NY, USA, 2006; ISBN 0791085627.

19. McMaster, G.S.; Green, T.R.; Erskine, R.H.; Edmunds, D.A.; Ascough, J.C. Spatial interrelationships between wheat phenology, thermal time, and terrain attributes. *Agron. J.* **2012**, *104*, 1110–1121.

20. Archontoulis, S. V.; Miguez, F.E. Nonlinear regression models and applications in agricultural research. *Agron. J.* **2015**, *107*, 786–798.

21. Chang, K.H.; Warland, J.S.; Bartlett, P.A.; Arain, A.M.; Yuan, F. A simple crop phenology algorithm in the land surface model CN-CLASS. *Agron. J.* **2014**, *106*, 297–308.

22. Yin, X.; Goudriaan, J.; Lantinga, E.A.; Vos, J.; Spiertz, H.J. A flexible sigmoid function of determinate growth. *Ann. Bot.* **2003**, *91*, 361–371.

23. YAN, W.; Hunt, L.A. An equation for modelling the temperature response of plants using only the cardinal temperatures. *Ann. Bot.* **1999**, 607–614.

24.   Patrignani, A.; Ochsner, T.E. Canopeo: A powerful new tool for measuring fractional green canopy cover. *Agron. J.* **2015**, *107*, 2312–2320.

25.   Mudereri, B.T.; Abdel-Rahman, E..; Dube, T.; Landmann, T.; Niassy, S.; Tonnang, H.E.Z.; Khan, Z.R. Potential of resampled multispectral data for detecting Desmodium-Brachiaria intercropped with maize in a "Push-Pull" system. *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *XLIII-B3-2*, 1017–1022.

26.   Gage, J.L.; Miller, N.D.; Spalding, E.P.; Kaeppler, S.M.; de Leon, N. TIPS: A system for automated image-based phenotyping of maize tassels. *Plant Methods* **2017**, *13*, 1–12.

27.   Crain, J.; Mondal, S.; Rutkoski, J.; Singh, R.P.; Poland, J. Combining High-Throughput Phenotyping and Genomic Information to Increase Prediction and Selection Accuracy in Wheat Breeding. *Plant Genome* **2018**, *11*, 1–14.

28.   Singh, P.; Boote, K.J.; Kadiyala, M.D.M.; Nedumaran, S.; Gupta, S.K.; Srinivas, K.; Bantilan, M.C.S. An assessment of yield gains under climate change due to genetic modification of pearl millet. *Sci. Total Environ.* **2017**, *601–602*, 1226–1237.

29.   Smale, M.; Byerlee, D.; Jayne, T. Maize revolutions in Sub-Saharan Africa. *An African Green Revolut. Find. Ways to Boost Product. Small Farms* 2013, *9789400757*, 165–195.

30.   Mudereri, B.T.; Abdel-Rahman, E.M.; Dube, T.; Landmann, T.; Khan, Z.; Kimathi, E.; Owino, R.; Niassy, S. Multi-source spatial data-based invasion risk modeling of Striga (Striga asiatica) in Zimbabwe. *GIScience Remote Sens.* **2020**, *57*, 553–571.

31.   Kiniry, J.R.; Bonhomme, R. Predicting maize phenology. *Predict. Crop Phenol.* **1991**, 115–131.

32.   Tonnang, H.E.Z.; Juarez, H; Carhuapoma, P.; Gonzales, J.C.; Mendoza, D.; Sporleder, M.; Simon, R.; Kroschel, J. ILCYM- Insect life cycle modeling : software for developing temperature-based insect phenology models with applications for regional and global pest risk assessments and mapping. *Int. potato centre. Lima Peru* **2013**.

33.   Tonnang, H.E.Z.; Hervé, B.D.B.; Biber-Freudenberger, L.; Salifu, D.; Subramanian, S.; Ngowi, V.B.; Guimapi, R.Y.A.; Anani, B.; Kakmeni, F.M.M.; Affognon, H.; et al. Advances in crop insect modelling methods—Towards a whole system approach. *Ecol. Modell.* **2017**, *354*, 88–103.

34.   Lampton, M. Damping-undamping strategies for the Levenberg-Marquardt nonlinear least-squares method. *Comput. Phys.* **1997**, *11*, 110–115.

35.   Steiniger, S.; Hunter, A.J. Free and Open Source GIS Software for Building a Spatial Data Infrastructure. *Geospatial Free Open Source Softw. 21st Century, Lect. Notes Geoinf. Cartogr.* **2012**, 247–261.

36.   R Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. 2020.

37.   Ripley, B.D. The R Project in Statistical Computing. *MSOR Connect.* **2001**, *1*, 23–25.

38.   Silva, V. Foundations of Eclipse RCP. *Pract. Eclipse Rich Client Platf. Proj.* 2009, 1–19.

39.   Elzhov, T.; Mullen, K.; Spiess, A.; Bolker, B. Package "minpack. lm." *Https://Cran.R-Project.Org/Web/Packages/Minpack.Lm/Minpack.Lm.Pdf* **2016**.

40.   Venables, W.N.; Ripley, B.D. Modern Applied Statistics with S. Fourth Edition. *Springer, New York. ISBN 0-387-95457-0* **2002**.

41.   Bivand, R.; Archer, E.; Baddeley, A.; Bearman, N.; Callahan, J.; Forrest, D.; Friendly, M.; Stokely, M. Package 'maptools' R topics documented : **2020**.

42.   Becker, R.A.; Wilks, A.R.; Brownrigg, R.; Minka, T.; Deckmyn, A.. maps: Draw Geographical Maps. R package version 3.3.0. **2018**, http://cran.r-project.org/package=maps.

43.   Urbanek, S. Rserve: A Fast way to provide R functionality to ppplications. *Proc. 3rd Int. Work. Distrib. Stat. Comput. Vienna, Austria* **2003**.

44.   Sharpe, P.J.H.; DeMichele, D.W. Reaction kinetics of poikilotherm development. *J. Theor. Biol.* **1977**, *64*, 649–670.

45.   Yin, X.; Kropff, M.J.; McLaren, G.; Visperas, R.M. A nonlinear model for crop development as a function of temperature. *Agric. For. Meteorol.* **1995**, *77*, 1–16.

46.   Rykiel, E.J.J. Testing ecological models: the meaning of validation. *Ecol. Modell.* **1996**, *90*, 229–244.

47.   Grassini, P.; Pittelkow, C.M.; Cassman, K.G.; Yang, H.S.; Archontoulis, S.; Licht, M.; Lamkey, K.R.; Ciampitti, I.A.; Coulter, J.A.; Brouder, S.M.; et al. Robust spatial frameworks for leveraging research on sustainable crop intensification. *Glob. Food Sec.* **2017**, *14*, 18–22.