

Spatial predictions of maize phenology: workflows and potential decision support applications

P. Craufurd, K. Oluoch, H. Tonnang¹
16-October-2020

1. Introduction

Choosing which variety (OPV or hybrid) to grow in a particular location is a very important decision for farmers as this determines to a very large extent the yield potential and the risk of crop failure. Many extension agencies ask which varieties are suitable or best adapted for different locations, as the list of released or approved varieties in any one country can be large. Furthermore, seed companies classify varieties for their suitability or adaptation in different ways, such as by altitude (e.g. suitable for lowland or <1000m) and duration (e.g. 120-150 d variety; long duration). While these may be indicative, they are not very precise or predictive, and often choosing the right variety comes down to trial and error.

The [TAMASA](#) project developed and piloted a decision-support tool called Maize Variety Selector (MVS) designed to provide advice on which of the many available varieties to plant in a given location, and the characteristics of different varieties. Based on experience in Ethiopia, Kenya, Nigeria and Tanzania with users and those breeding or releasing varieties, we developed a simpler workflow to generate the predictions needed for decision-support or advice. The steps are:

- (i) generate the data needed to parameterize or model the response of different varieties or genotypes to temperature, used to;
- (ii) generate model calibrations (parameterization) that enable times of flowering and maturity (harvest) of varieties to be predicted;
- (iii) download spatial or gridded temperature (weather) databases to make and map predictions;
- (iv) use tools that use the calibrated models to generate a database of spatial predictions of flowering and maturity at any location and any sowing date; and
- (v) map predictions spatially.

This workflow will not result in an app *per se*, but the resulting spatial database may serve as the basis of an app or any other decision tool.

The rest of this note is structured as follows. First, we describe why adaptation is important ([section 2](#)) and how the response to temperature can be modelled ([section 3](#)). Then we go through the five steps of the workflow ([section 4](#)). Lastly, we outline a decision-support system for providing advice on which varieties to plant where ([section 5](#)).

¹ Peter Craufurd (p.craufurd@cgiar.org), Kevin Oluoch (kevin.o.oluoch@gmail.com), Henri Tonnang (htonnang@icipe.org).

2. Why is choosing the right duration or maturity important?

Sowing date and variety duration or type, are decisions farmers have some choice over to improve adaptation and to manage harvest dates. Growing varieties of any crop species that are well adapted to a particular location is very important for productivity. In the rainfed environments in which maize is predominantly grown in SSA, adaptation is largely determined by the timing of sowing, flowering and maturity in relation to rainfall and available soil moisture (Fig. 1). If a variety flowers and matures too late, yield will be reduced by water, and in some cases heat, stress. If a variety flowers and matures too early, then yield will be below the maximum or potential because the growing season was shorter than it need be and all the available water was not exploited. The timing of flowering, the most sensitive stage of crop development, is also very important for minimizing pest and disease damage. It is worth noting here that many traditional or landrace varieties of crops such as sorghum or pigeon pea are sensitive to photoperiod or day-length, a mechanism that ensures timely flowering by adjusting the length of the vegetative period (see Craufurd & Qi, 2001 for an example).

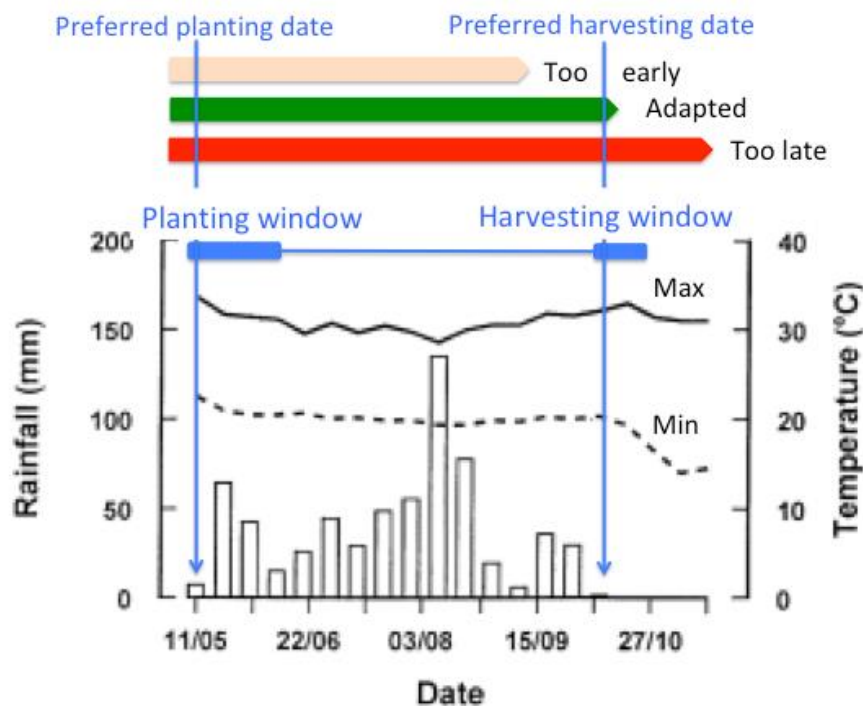


Figure 1. An exemplar of a typical mono-modal growing season (Zaria in Nigeria) showing the rainfall pattern, preferred planting and harvesting dates, and the duration of three hypothetical varieties. Based on Craufurd and Qi (2001).

3. How is (maize) duration (phenology) modelled?

Phenology or development of maize from sowing through flowering (tasseling and/or silking) to maturity (black layer formation) is governed by a well-defined response to temperature, usually expressed as a rate or the reciprocal of duration (Figure 2). The response is typically described in terms of the so-called cardinal temperatures, namely the base (T_{\min} or T_{base}), optimum (T_{opt}) or maximum (T_{\max}) temperature, and constant(s) describing the rate response. So, at cool temperatures ($\sim 10^{\circ}\text{C}$ in most tropical crops) rate of development is effectively zero. As the temperature increases between the base and optimum, the rate of development increases and flowering, for example, occurs sooner at warmer than cooler temperatures, and earliest at the optimum (which effectively has a range of a few degrees). The optimum temperature is typically around 30° to 34°C . At higher temperatures the rate of development is reduced and flowering is delayed.

The response to temperature of individual varieties can be quantified from experimental data, either from multi-location trials or from designed calibration experiments, and models or curves fitted to the observed data relating rate of development to average (daily) temperature used to predict flowering and maturity.

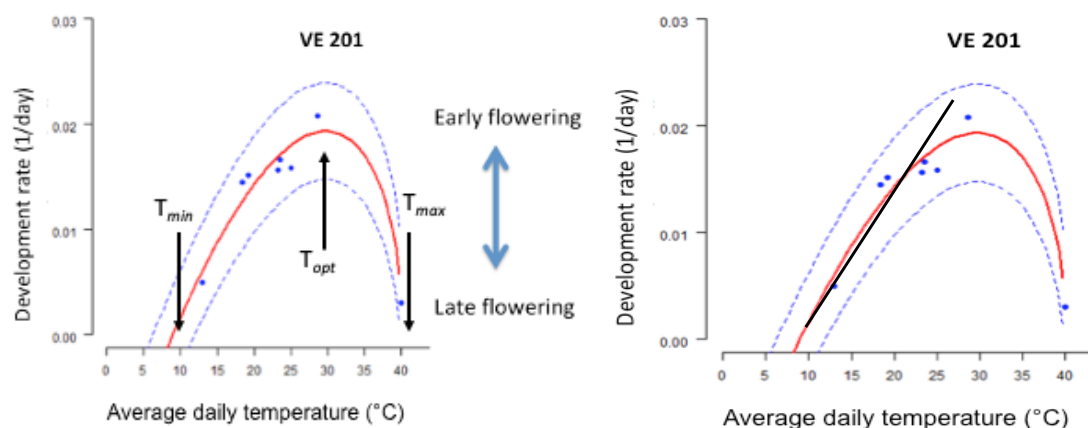


Figure 2. The response of duration to flowering of maize variety VE201 to temperature.

Note that the response is modeled as the rate of development (1/duration or the reciprocal of duration). The figure on the left has a polynomial fitted and the one on the right has a linear regression superimposed between the minimum and optimum. The minimum or base (T_{\min}), optimum (T_{opt}) and maximum (T_{\max}) temperature are also shown.

Source: Tonnang et al. (2018)

In most cases, simple accumulated thermal time or degree-day ($^{\circ}\text{C day}$) targets above a defined minimum or base temperature (typically around 10°C in maize) can be estimated by assuming a linear response to average temperature. This is because in most growing seasons in most environments the optimum temperature will not be exceeded or not be exceeded for many days. Therefore,

in Figure 2, if the regression line where rate = $a + b * T$ (average daily temperature) is given by:

$$\text{Rate or } 1/\text{days to flowering} = 0.013 + 0.00133T$$

The minimum or base temperature is given by:

$$T_{\min} = a/b \text{ or } 0.013/0.0013, \text{ and hence } T_{\min} = 10^{\circ}\text{C}.$$

The accumulated thermal time or degree days to flowering is given by the reciprocal of b :

$$1/b \text{ or } 1/0.00133, \text{ and hence } 750^{\circ}\text{C days}.$$

Thermal time simulations simply sum the difference between average daily temperature and T_{\min} :

$$\text{Degree days } (^{\circ}\text{Cd}) = \sum T - T_{\min}$$

until the target of 750°C days is achieved. Thus, if the average daily temperature was 25°C, then each day 25 minus 10°C or 15°Cd would accumulate, and flowering would take 750/15 or 50 days.

Where data includes high temperatures, the thermal time approach can be modified to account for the negative or delaying effect of high temperature, where if $T > 34^{\circ}\text{C}$, then:

$$\text{Degree days } (^{\circ}\text{Cd}) = (T_{\text{opt}} - T_{\min}) * (1 - (T - T_{\min}) / (T_{\max} - T_{\text{opt}})).$$

The same principle applies if rates rather than thermal times are modelled; each day a fraction of progress (from zero to one) to flowering or maturity occurs or accumulates until the target of one is reached.

4. Workflow

4.1. How can maize flowering and/or maturity data for calibration be collected and operationalized?

Temperature-dependent models for maize (and other crops) can be calibrated from two sources of data, namely multi-locational trial data or calibration experiments. It is recommended to start with calibration experiments if possible because these: (a) allow one to include all the varieties of interest (i.e. not just breeding lines or new releases) and (b) choose appropriate sites for observing varieties.

4.1.1 Calibration experiments

Calibration experiments or trials are designed based on the temperature response shown in Fig. 2. In principle, and practice, temperature response curves

or models can be fitted from relatively few observations of flowering or maturity as long as these data represent/cover as much variation or range in average temperature as possible. A minimum of four experimental sites should be selected with low (cool) to warm (optimum), and hot (above optimum) average temperatures to observe flowering and/or maturity. As long as the range in average temperature is large enough, as in Fig. 2 where observations were made at about 12°, 17°, 22° and 25°C – and ideally one warmer temperature, the response curve can be adequately fitted. It is possible to do these calibration experiments in controlled environments (see Erskine et al. for an example), but usually there are too many varieties that need to be calibrated for this to be practical (and assuming such facilities are available of course). Some examples of calibration trials used in TAMASA and others are given below.

A few key points for calibration trials:

- Observing flowering and maturity, both which should be defined based on standard descriptors (https://cropgenebank.sgrp.cgiar.org/images/file/maize/Key_Access_Descriptors_Maize.pdf), provides the most useful information. However, just observing maturity would also suffice as it is crop duration (harvest date) that farmers are most interested in.
- Calibration trials should be sited across an altitude gradient, which is an easy way to create the required variation in average temperature (see the examples below).
- Including more than one sowing date will also create some variation in growing temperature (and provide extra insurance or replication).
- Daily maximum and minimum temperature (or more frequent intervals with digital devices) must be measured at or very near each site. These temperature data are essential for modelling or curve fitting. Most research stations do have weather stations so are logical choices as sites.
- The calibration trials can be very simple and replicated experiments are not needed. A single replicate of a single 2 m row (for maize) is all that is needed to observe flowering and maturity. Development, as opposed to growth, is less sensitive to growing conditions and it is usually only when stress is severe that the timing for flowering, for example, is affected. So, a single replicate is good enough. In this way many varieties can be evaluated at one time without using much space.
- In general, most research and breeding programs will record when 50% of the plants in a row have flowered or are mature. This is a simple observation.

Examples

- Tonnang et al. (2020) used altitude to create different temperature environments for characterizing varieties in Ethiopia and Tanzania.

Table 1. Study sites in Tonnang (2020)

Site, Country	Altitude (m)	Ave. min. (°C)	Ave. max. (°C)	Days to flowering
---------------	--------------	----------------	----------------	-------------------

Dedessa, ET	1224	18.5	32.2	65
Bako, ET	1640	14.1	27.2	80
Holleta, ET	2351	8.8	22.4	112
Miwaleni, TZ	1581	20.0	31.5	66
Uyole, TZ	1769	12.7	25.4	82
Igeri, TZ	2212	11.4	20.3	103

- Craufurd et al. (1999) characterised sorghum for responses to temperature and photoperiod using an altitude gradient in Kenya. In this study sorghum germplasm was grown at a range of sites in Kenya from sea level to >1600m to create variation in average temperature. At some sites daylength was also extended to simulate higher latitudes.
- Erskine et al. (1990) characterised lentil germplasm for photoperiod and temperature responsiveness using four controlled environments. In this study lentil germplasm was grown in four controlled environments using two temperatures and two photoperiods.

4.1.2 Using data from multi-locational variety release trials.

Dates of flowering and maturity are routinely recorded in breeder and variety release trials (NPTs or similar), and often these are on research stations where daily weather data are also recorded. Potentially these data can be used to model flowering or maturity, especially if trial sites capture a good range in average temperature, as outlined above. The data in Figure 2 is from CIMMYT multi-locational breeder trials (Tonnang et al., 2018); see also Adnan et al. (2019) who compared parameterization and predictions performed using a calibration data set and data from breeder's trials in Nigeria. The study of Adnan et al. (2019) in particular showed that models derived from a small subset (number) of trials in northern Nigeria was able to generate robust predictions similar to those from a larger, designed set of trials.

The best strategy would be to plant all the available varieties (released and local, i.e. what is actually grown by farmers and therefore the varieties they want information on) in calibration trials to quickly generate the data needed for models and predictions. Subsequently, as new varieties are released, the NPT data should be sufficient to provide the necessary calibration.

4.2. Parameterizing curves/models of flowering and maturity

Time to, or the date of, flowering and maturity can be modeled (i.e. Fig. 2, the equations describing the response to temperature parameterized) in a number of applications, including for example GENCALC or DSSAT (see Adnan et al. 2019). Scripts to fit curves can also be written in most programs, including R.

TAMASA developed an open source software in R, called PPMaP, described below.

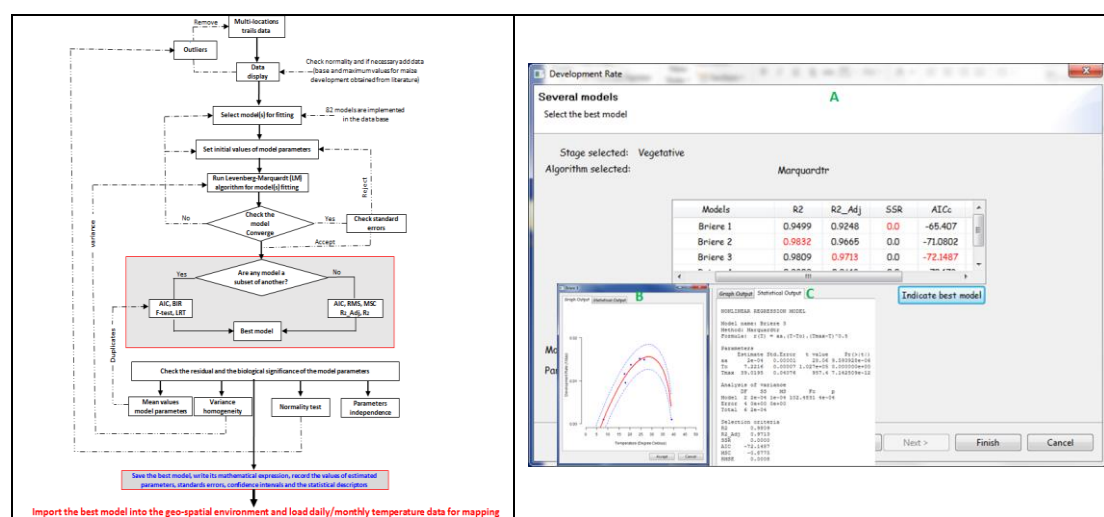
To run any model to optimize and parameterize the equations describing the curve(s) requires:

- the observed dates or times (days after sowing or emergence) for flowering or maturity;
- the average daily temperature for the season/ year or for the duration of the experiment.
- photoperiod or daylength: Note that in the case of maize all the varieties we have tested have been insensitive to photoperiod or daylength and hence only require temperature to make predictions. Some varieties may respond to photoperiod and require a parameter for this as well.

There is, unfortunately no standard equation used to model phenology. A simple linear or thermal time response (described above) will suffice for most practical purposes (i.e. to identify which varieties are relatively early or late). If the range of temperatures used to parameterize the model or curve was small, then assuming a constant T_{\min} (for example 8°C as in CERES/DSSAT), and if needed a constant T_{\max} (40°C), is justified.

PPMaP

Plant Phenology Prediction and Mapping (Tonnang et al., 2020) is an open-source software for modelling response to temperature (there are 82 models coded in the application) and mapping the predictions. The software is written in R with an easy to use GUI interface. The software uses observed phenology and temperature data to assess and fit models. The paper by Tonnang et al. (2018) describes the development of the software and an example from Kenya. The mapping component is described later.



The code for PPMaP and a user manual can be found here <https://github.com/Atoundem/PPMaP>. Basically, the user has to enter: (i) values of the rate of development (1/duration to flowering or maturity); and (ii) the average daily temperature over the same period from different experiments or

trials, and the software will fit and determine the best model. The resultant model can be used for mapping in PPMaP itself or in the XLS tool, both described below.

GENCALC

The GENCALC program of the DSSAT (Version 4.6) can be used to calibrate the parameters or coefficients used by DSSAT and CERES models to predict phenology (see Adnan et al. 2019 for an example). Four parameters define the life cycle and are needed to simulate flowering and maturity (Boote et al., 2003):

- P1 (° days): Thermal time from seedling emergence to the end of juvenile phase (taken as panicle initiation)
- P2 (° days): Delay in development for each hour that day-length is above 12.5 hours
- P5 (° days). Thermal time from silking to time of physiological maturity
- PHINT (°C day tip⁻¹): Thermal time between successive leaf tip appearance

These model parameters are all basically thermal time or degree-day values and hence easy to use. Ideally panicle initiation, silking, physiological maturity and leaf-tip appearance would be observed, though GENCALC will optimize some parameters based on initial values. Most maize varieties are not affected by photoperiod and P2 can be set to a small value (Adnan et al. used P2=0.5 for maize in northern Nigeria, for example). Some typical values from that study are given in the table below.

Table 2. Parameters to predict flowering and maturity date in selected maize varieties from Nigeria (Adnan et al., 2019)

Variety	P1 (°C days)	P5 (°C days)	PHINT (°C day tip ⁻¹)
Ife hybrid 6	224	521	36.9
Sammaz 42	230	683	45.5
OBA SUPER I	293	768	45.0
SC 8325	290	782	41.2
Sammaz 32	282	601	45.0

4.3. Downloading spatial temperature (weather) databases

There are a number of sources of daily, gridded weather data suitable for making spatial predictions. Schauburger et al. (2020) provide a current list of all the available databases. Daily data is preferred to monthly data to capture any seasonal variation in temperature.

For the XLS-based macro described below (section 4.4: MSPP) we used AgMERRA data (<https://data.giss.nasa.gov/impacts/agmipcf/>), which is 25 km (~0.25°) resolution, gridded dataset used by the crop modeling community. AgMERRA has daily weather data from 1980 to 2010 and is free to download in a range of formats, including CSV. A 25 km² resolution pixel is rather coarse for

areas with a large variation in topography, but the prediction is primarily to allow a comparison of varieties rather than make an exact prediction.

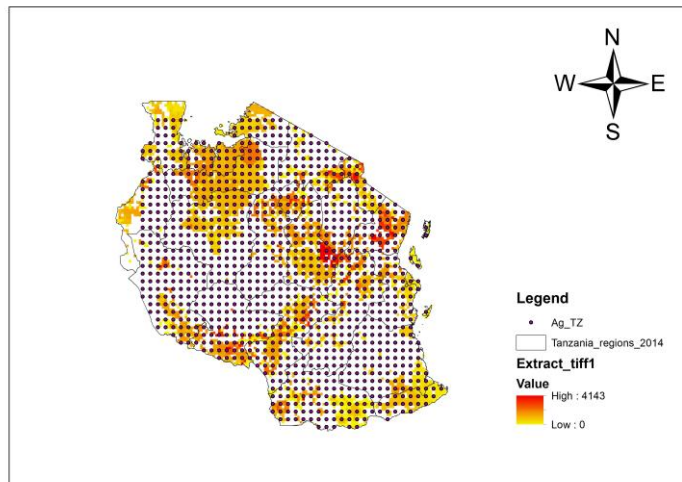


Figure 3. AgMERRA grids for Tanzania (~480 grids) overlaid on the cropland area (~180 grids)

The R file `agMERRAnc4ToCSV` (<https://github.com/KevinOluoch/MSPP>) provides the code needed to use the AgMERRA netCDF format files downloaded from the site and convert them into CSV files that can be used in other applications.

Create data for the MSPP

The temperature data used in MSPP is available on the agMERRA website (<https://data.giss.nasa.gov/impacts/agmipcf/>) in the netCDF format. The spatial extent of the data is the whole world. Each NetCDF file covers a wide range of climate variables but we are only interested in either the minimum or maximum daily temperature, or the average daily temperature over the period between 1980 and 2010. The first step is thus to download these data for all the years (or a required period) and have the data in (three folders): minimum, maximum and average temperature. The file paths to these folders are then used to get the file paths of all the NetCDF files in them:

```
library(devtools)

## Loading required package: usethis

# Load "agMERRAnc4ToCSV" package
path2package <- "../Rpackage/agMERRAnc4ToCSV"
devtools::load_all(path2package)

## Loading agMERRAnc4ToCSV

# Place inputs in the data folder
netCDF_TminFolder <- "../data/netCDF/AgMERRA_tmin"
netCDF_TmaxFolder <- "../data/netCDF/AgMERRA_tmax"
netCDF_TavgFolder <- "../data/netCDF/AgMERRA_tavg"
```

```
# Get netCDF filepaths
Tmin.nc4files <- agMERRAnc4ToCSV::nc4fullpaths(netCDF_TminFolder)
Tmax.nc4files <- agMERRAnc4ToCSV::nc4fullpaths(netCDF_TmaxFolder)
Tavg.nc4files <- agMERRAnc4ToCSV::nc4fullpaths(netCDF_TavgFolder)
```

MSPP uses point data in its computation. To use it in a given set of locations or geographical area, you need to extract the data. You need to have the location data as a “SpatialPointsDataFrame” object or in a CSV file, which can be converted to a “SpatialPointsDataFrame” using the “wgsSpatialPointsDataFrame” function. This function assumes that the coordinates are in WGS84 if no coordinate system is provided:

```
# Provide a path to the Locations CSV
inputCSV <- "../data/tzaLocations.csv"

# Create a SpatialPointsDataFrame Object from the CSV
shpfile <- agMERRAnc4ToCSV::wgsSpatialPointsDataFrame(inputCSV,
                                                       LongitudeColumn = "Lon",
                                                       LatitudeColumn = "Lat")
```

Extracting data from the NetCDF files and providing the output in a data frame involves a few steps which are all covered in the “xtrctagMERRAnc4” function from “agMERRAnc4ToCSV” package. The “xtrctagMERRAnc4” function requires a vector of file paths to NetCDF files, a SpatialPointsDataFrame of locations to be extracted and the names of two columns in the SpatialPointsDataFrame with geographical names. It then returns a data frame of extracted values, which is organized in the format required in the MSPP tool.

The data should be saved in CSV files that can either be pasted in the MSPP tool or loaded via links to Excel sheets. The steps for connecting an Excel sheet to an CSV file is provided below:

```
# Write to CSV
dir.create("temperatureCSVs", showWarnings = FALSE)
write.csv(Tmin, "temperatureCSVs/Tmin.csv", row.names = FALSE, col.names = FALSE)
write.csv(Tmax, "temperatureCSVs/Tmax.csv", row.names = FALSE, col.names = FALSE)
write.csv(Tavg, "temperatureCSVs/Tavg.csv", row.names = FALSE, col.names = FALSE)
```

4.4. Tools to generate spatial predictions of maize phenology

To recap, the duration from sowing to flowering and/or maturity in maize is determined by average daily temperature between minimum and optimum cardinal values. Gridded (i.e. geo-referenced data linked to a pixel or cell of a

certain size) daily temperature data for individual years or long-term averages are available across the globe from several sources at resolutions of 50 km (0.5 deg) to 10 km (0.1 deg). Therefore, given a gridded (spatial) temperature database, and model equations defining the response to temperature, the duration from sowing to flowering and maturity can be predicted for any given sowing date at each grid point or pixel.

Below we give two examples for generating predictions, an XLS-based tool MSPP (see below and 4.5.1), and an open-source R/Java software called PPMaP (Tonnang et al., 2020; see 4.5.2).

MSPP XLS-tool

The MSPP tool is a relatively simple piece of software that runs in Excel to generate a database of predictions (available here: <https://github.com/KevinOluoch/MSPP>). MSPP is written in Visual Basic (VBA) programming language that is accessible in Excel via the developer tab. This software should therefore be accessible and useable to most people familiar with Excel and does not require any other operating languages such as R or Python.

There are five steps to generate the predictions using this tool once the user has the model parameters (Step II) and the AgMERRA database (Step III):

1. Optional. The AgMERRA database will have all the grid cells in a country, for example 480 or so for Tanzania. However, many of these cells will not contain crops and are therefore not needed for the predictions. Removing these as a first step can reduce the number of simulations and the size of the final database of predictions, which is helpful. A cropland mask (e.g. global croplands: <https://croplands.org/app/map?lat=0.17578&lng=0&zoom=2>) can be used to remove all the pixels or grid cells where crops were not grown;
2. The AgMERRA database is first imported in to the MSPP tool; see Section III above
3. The user then chooses the appropriate model or equation and enters the parameters from the calibration. Users can simulate sowing to flowering or maturity, or if flowering and maturity have been calibrated separately, simulate sowing to flowering, and flowering to maturity.
4. The user then chooses sowing date(s) (1, 10 or 20th day of the month or dekads), month(s) and year(s) for the simulations. Three sowing dates per month is adequate for comparing varieties across sowing dates and providing advisories.
5. The tool predicts flowering and/or maturity in every pixel for each selected sowing date, month and year and outputs these as an XLS file. If multiple years were used, then the tool also calculated the average plus the standard deviation and generates an XLS output file. These files can be used in other applications.

The MSPP tool can be found here (<https://github.com/KevinOluoch/MSPP>) and instructions are given on each tab/screen when the file is opened.

4.5. Mapping spatial predictions

Once the predictions are made, they are saved into an Excel workbook (database) that allows any predicted value of flowering and/or maturity to be extracted for a given sowing date. These values can also be mapped.

The MSPP tool's output is an Excel workbook. The workbook has two or three worksheets, depending on whether flowering and/or maturity was predicted. There is a summary sheet, which gives the mean and standard deviation of the number the predictions over the 30 years (or selected period) that MSPP was run at. The other sheets give the predictions for each sowing date in all selected months and all selected years. The summary data are the most useful ones to use for mapping predictions or as an app database.

4.5.1 Mapping predictions from MSPP

Interpolating the Mean values from the summary sheet can be done using the "interpolateMeanDays" function (agMERRAnc4ToCSV library). The function takes the coordinates of the locations and the mean values as input parameters and returns a raster of the interpolated values over an area covering the extent of the input coordinates:

```
data.df <- xlsx::read.xlsx("../data/Briere_3_1980_2009_02_01.xlsx",
1)

Flowering.raster <- agMERRAnc4ToCSV::interpolateMeanDays(Lon = data.
df$Lon,
                                                         Lat = data.df$Lat,
                                                         meandays = data.df$Flowe
ring.Mean)
## [inverse distance weighted interpolation]

Maturity.raster <- agMERRAnc4ToCSV::interpolateMeanDays(Lon = data.d
f$Lon,
                                                         Lat = data.df$Lat,
                                                         meandays = data.df$Matur
ity.Mean)
## [inverse distance weighted interpolation]

plot(Flowering.raster,
      main="Sowing to Flowering Mean Days",
      xlab="Longitude",
      ylab="Latitude")
```

4.5.2 Predictions and mapping in PPMaP

The PPMaP software (<https://github.com/Atoundem/PPMaP>) will also map the phenology predictions if a suitable database is loaded into PPMaP. The User Manual explains how to do this.

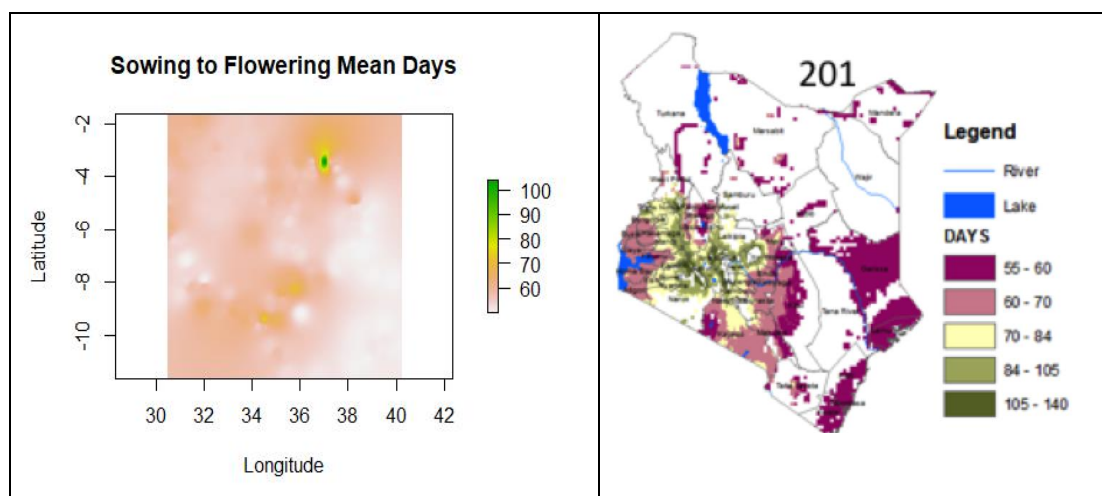


Figure 4. Maps generated by interpolateMeanDays R code (left) and PPMaP (right).

5. Suggested application or decision-support system for advisories on variety choice

The workflow or logic (Fig 5) below is an example of a simple app or decision-support system based on the outputs of the MSPP tool. This could be implemented through ODK (Open Data Kit) or of course made into an Android app. This is based on TAMASA's experience with an application called MVS that was tested with users in Nigeria, Ethiopia and Tanzania.

The opening screen will describe what MVS does and how it is used. The primary purpose of MVS is to help identify which varieties 'fit' the cropping calendar of the beneficiary, usually a farmer. The tool can be used by service providers, who are in our experience likely to be the primary users, or by individual farmers.

The first screen will allow the user to identify the location where they want to plant maize. This will be through nested region/zone-state-district to lowest administrative unit. This is simpler to implement (at least in ODK) than clicking on a map. We also observed that with a map function the user spent a lot of time looking for their village name, which is often not marked or the name not easily visible. A nested site schema also makes it easy to link the relevant pixel with the predictions.

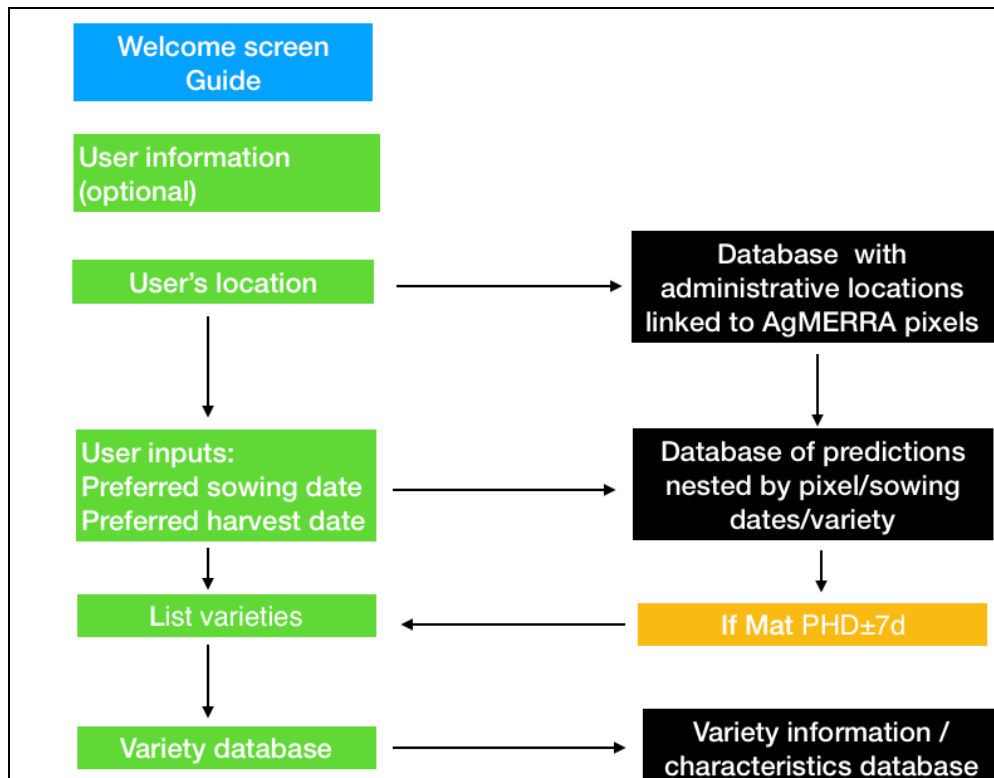


Figure 5. Logic of MVS with user inputs and app outputs

The second screen will ask users to input a preferred sowing or planting date and a preferred harvest date. This is predicated on the assumption that farmers or other users know their own growing or cropping season well and can give realistic start and end dates. Farmers usually know when they want to harvest and, in many cases, they want to spread the date of harvest if possible. This logic also keeps the application fairly simple in that it only requires two inputs from farmers; preferred sowing date and preferred harvest date. Figure 1, at the start of this document, explains the logic of this decision.

The application would then extract the flowering and/or maturity date for every variety at the given sowing date in the relevant pixel, and display only the varieties that mature within, say, ± 7 d of the user preferred harvest date. This criterion can be varied and/or determined for each country, and of course all predictions could be shown if desired.

Users can also click on the varieties listed and see the available information (characteristics) on those varieties. We have observed in a number of applications that farmers valued information about varieties. There will also be an option to try another sowing date or quit.

Information on variety characteristics and suitability needs to be linked to information on where seed can be purchased to improve the utility of MVS. A good additional option would be to include an agro-dealer database with an option to show the nearest agro-dealer and give a phone number. In Nigeria TAMASA did map agro-dealers and seed dealers and make a static database.

An example of a Decision-Support System for maize varieties in Mbeya District, Southern Zone, Tanzania

Mbeya District, comprising 10 sub-Districts centered around the city of Mbeya (Fig. 6), is a major maize growing area in the Southern Highlands of Tanzania. This District includes highland and lower altitude (valley) areas with different climates and ecologies that could influence the choice of variety for different locations.

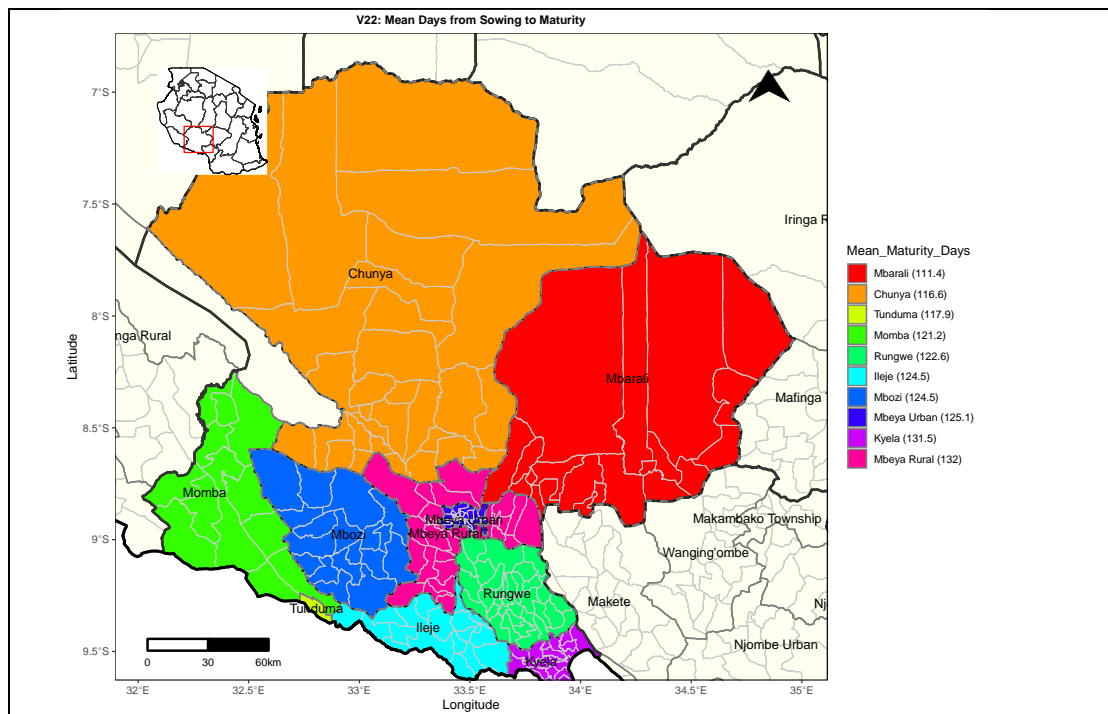


Figure 6. Mbeya District and sub-Districts.

More than 30 widely grown maize varieties and hybrids were grown in experiments in Tanzania and their duration from sowing to maturity modelled. These models allow for the time of date of maturity to be predicted for any given sowing date at any given location. We can answer a number of practical questions from these data.

How does time to maturity vary with sowing date?

The duration from sowing to maturity in maize is affected by temperature, which can vary during the season (and of course between sites, especially if the altitude is different). In Chunya late sowings do take about 15 days longer to mature than early sowings, due to cooler temperatures in the spring (Fig. 7). However, early sowings are not much affected.

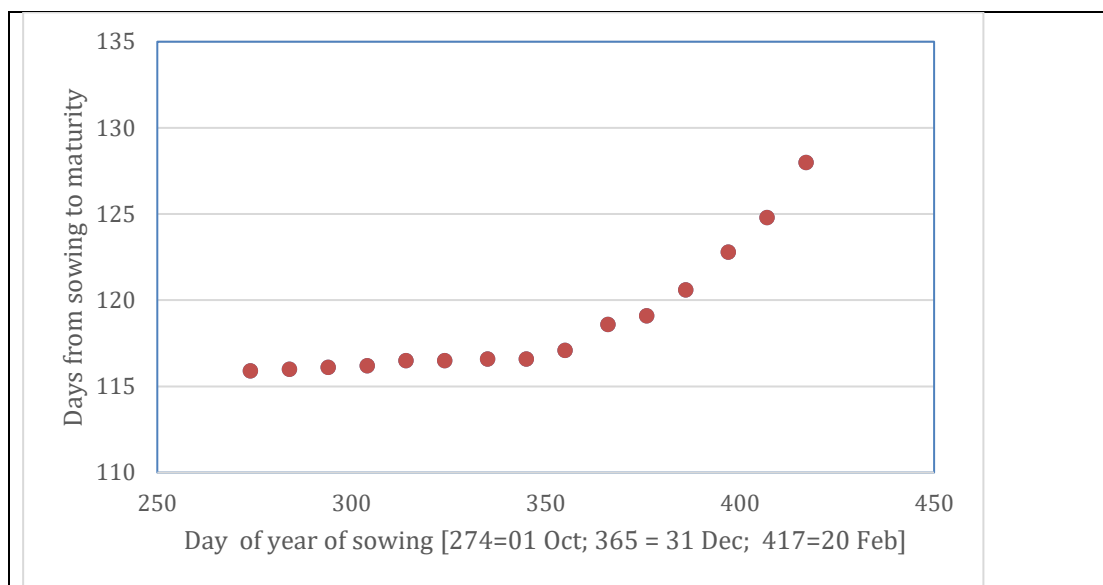


Figure 7. Predicted response of a medium variety to sowing date at Chunya sub-District

If I plant on 01 December in Chunya sub-District, when will different varieties be ready for harvest?

Taking a typical sowing date in the region of 01 December, if we compare varieties at Kiwanja in Chunga sub-District as an example, we can see that most varieties matured between 100 and 120 days after sowing, i.e. between 01 and 20 March (Fig. 8). One variety was very early, maturing in less than 90 days (~15 February) and a few varieties were very late, maturing in more than 130 days (~01 April). Overall, these varieties offer growers substantial variation in maturity date to choose from. An app or paper-based decision-support tool would identify those varieties predicted to mature in a period closest to that desired by the grower.

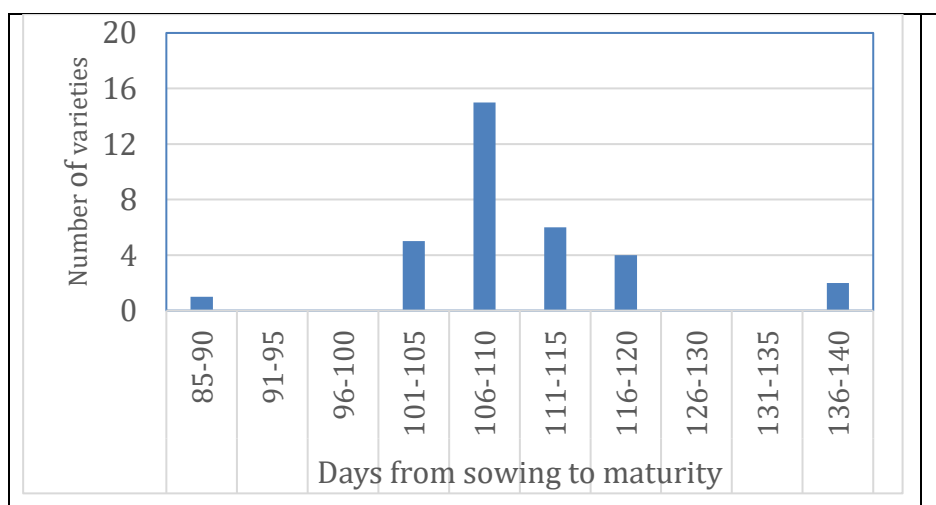


Figure 8. Predicted variation in time from sowing to maturity among varieties in Chunya, Tanzania

If I want to advise growers in Mbeya District on which variety to grow, how can I do this?

The two maps below (Fig. 9) show the spatial patterns for maturity date of an early and medium variety across Mbeya District for a 01 December sowing date. We can see from the darker colours where the highland and cooler areas are and that these result in more than 20 days difference in maturity within a variety. We can also see that in most sub-Districts there is 10 to 15 days variation in maturity date within an individual variety. While variety maturity could be mapped by District, as is done in Fig. 6 for the medium variety, it is better to delineate zones within sub-Districts for recommendations.

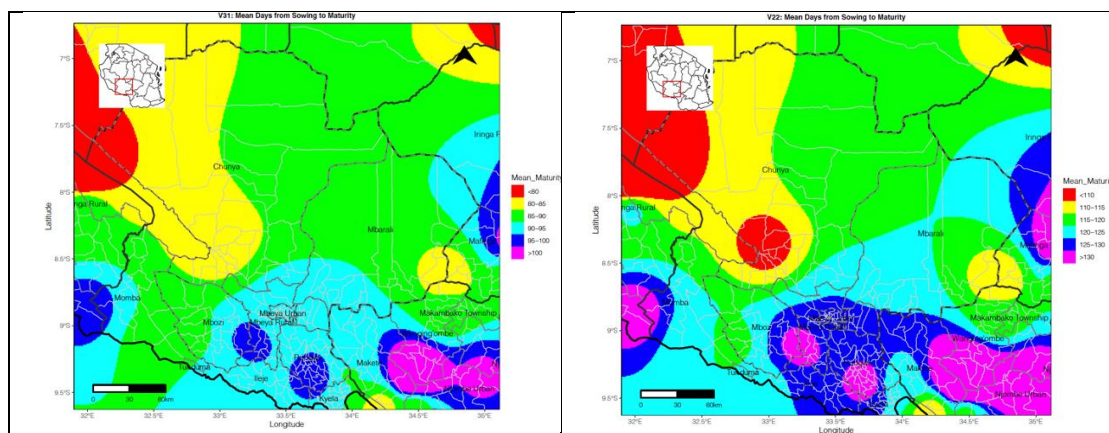


Figure 9. Predicted maturity dates for maize varieties V31 (early maturity) and V22 (medium maturity) in Mbeya District for a 01 December planting date

At the level of a District extension office, or for the AoI of an NGO, maps such as those in Fig. 9 could be generated to show farmers the relative differences in variety maturity for different sowing dates to aid decision-making. Tables or databases by town or community could also be produced with the same information. As these are based on long-term weather data, they are essentially a reference resource and do not need re-generating each year.

Code and software resources

MSPP and AgMERRA code : <https://github.com/KevinOluoch/MSPP>
PPMaP: <https://github.com/Atoundem/PPMaP>)

References

Adnan, A. A., Diels, J., Jibrin, J. M., Kamara, A. Y., Craufurd, P., Shaibu, A. S., ... & Tonnang, H. E. (2019). Options for calibrating CERES-maize genotype specific parameters under data-scarce environments. *PloS ONE*, 14(2), e0200118. <https://doi.org/10.1371/journal.pone.0200118>

- Craufurd, P. Q., & Qi, A. (2001). Photothermal adaptation of sorghum (*Sorghum bicolor*) in Nigeria. *Agricultural and Forest Meteorology*, 108(3), 199-211. [https://doi.org/10.1016/S0168-1923\(01\)00241-6](https://doi.org/10.1016/S0168-1923(01)00241-6)
- Craufurd, P. Q., Mahalakshmi, V., Bidinger, F. R., Mukuru, S. Z., Chantereau, J., Omanga, P. A., ... & Hammer, G. L. (1999). Adaptation of sorghum: characterisation of genotypic flowering responses to temperature and photoperiod. *Theoretical and Applied Genetics*, 99(5), 900-911. <https://doi.org/10.1007/s001220051311>
- Erskine, W., Ellis, R. H., Summerfield, R. J., Roberts, E. H., & Hussain, A. (1990). Characterization of responses to temperature and photoperiod for time to flowering in a world lentil collection. *Theoretical and Applied Genetics*, 80(2), 193-199. <https://doi.org/10.1007/BF00224386>
- Schauberger, B., Jägermeyr, J., & Gornott, C. (2020). A systematic review of local to regional yield forecasting approaches and frequently used data resources. *European Journal of Agronomy*, 120, 126153. <https://doi.org/10.1016/j.eja.2020.126153>
- Tonnang, H. E., Balemi, T., Masuki, K. F., Mohammed, I., Adewopo, J., Adnan, A. A., ... & Craufurd, P. (2020). Rapid acquisition, management, and analysis of spatial Maize (*Zea mays* L.) phenological data—Towards ‘Big Data’ for agronomy transformation in Africa. *Agronomy*, 10(9), 1363. <https://doi.org/10.3390/agronomy10091363>
- Tonnang, H. E., Makumbi, D., & Craufurd, P. (2018). Methodological approach for predicting and mapping the phenological adaptation of tropical maize (*Zea mays* L.) using multi-environment trials. *Plant Methods*, 14(1), 108. <https://doi.org/10.1186/s13007-018-0375-7>
- Tonnang, H. E., Guimapi, R. A., Anani, B., Makumbi, D., Mudereri, B., Balemi, T., & Craufurd, P. (2020). PPMaP: reproducible and extensible open-source software for plant phenological phase duration prediction and mapping. Preprints 2020, 2020090563 <https://doi.org/10.20944/preprints202009.0563.v1>