CSIRO AGRICULTURE AND FOOD
www.csiro.au

Re-evaluating the Prediction Models for Cotton Development

Michael Bange, Nic Finger, Jane Caton (CSIRO Agriculture and Food)

James Mahan, Paxton Payton (USDA ARS)

Nov 2017

A report for prepared for the Cotton Research and Development Corporation

Citation

Bange M, Mahan J, Payton P, Finger N, and Caton S. (2017) Re-evaluating the Prediction Models for Cotton Development. CSIRO and CRDC, Australia.

Copyright

Important disclaimer

Summary

Key management recommendations rely on accurate estimates of crop development a using the day degree approach. The day degree approach is a fundamental tool used to assess crop development against growth and management (e.g. nutrition sampling, first irrigation) milestones for that particular season's climate. Currently, the 'day degree' approach (Constable and Shaw, 1988; Bange and Milroy, 2004) is not robust to accommodate extremes of climate (heat/cold). There is a need to refine this approach to ensure the accuracy of this critical tool to accommodate temperature extremes and ensure we can use it confidently for management decisions in new cotton regions (e.g. Griffith). New approaches will be developed to accommodate temperature extremes thus improving predictive capabilities and management recommendations that rely on this approach.

To validate new day degree modelling approaches developed, crop development data (first square and flower) were recorded across the industry. Industry outcomes includes an alternative approach using day degrees that delivers more precise predictions and assessments of crop development for all cotton regions that will facilitate more accurate growth assessment and management decisions. Science outcomes include a published alternative methodology on cotton's crop development in response to wider temperature extremes.

This study was able to demonstrate that there were improvements in the predictability of time of first square and first flower measured in cotton crops. Two functions were able to better predict these phenological stages compared to the existing function used currently in the Australian industry (Constable and Shaw, 1988). The best performing functions were a variable temperature day degree function that used a base temperature of 15 °C and an optimum of 32 °C, and a physiological rate function that reflected similar temperature characteristics as the variable function. The use of these functions should be considered in the development of new cotton crop predictive capabilities as they will be able to account for more temperature extremes (high and low, that maybe more prevalent in a changing climate) and where cotton production moves into new regions.

The analyses of functions here also support the use of a base temperature of 15 °C (60 °F) used in USA cotton systems

Acknowledgments

Project support provided by the Cotton Research and Development Corporation, the CSIRO, and Cotton Inc. USA.

Jane Caton, Nic Finger and Darin Hodgson for collection of infield data from CSIRO experiments.

The CSD E&D team for collection of in-field data.

USDA Lubbock technical team, Texas for collection of in-field data from the USA.

Introduction

Key management recommendations rely on accurate estimates of crop development a using the day degree approach. The day degree approach is a fundamental tool used to assess crop development against growth and management (e.g. nutrition sampling, first irrigation) milestones for that particular season's climate. Currently, the 'day degree' approach (Constable and Shaw, 1988; Bange and Milroy, 2004) is not robust to accommodate extremes of climate (heat/cold). There is a need to refine this approach to ensure the accuracy of this critical tool to accommodate temperature extremes and ensure we can use it confidently for management decisions in new cotton regions (e.g. Griffith). New approaches will be developed to accommodate temperature extremes thus improving predictive capabilities and management recommendations that rely on this approach.

To validate new day degree modelling approaches developed, crop development data (first square and flower) were recorded across the industry. Industry outcomes includes an alternative approach using day degrees that delivers more precise predictions and assessments of crop development for all cotton regions that will facilitate more accurate growth assessment and management decisions. Science outcomes include a published alternative methodology on cotton's crop development in response to wider temperature extremes.

Methodology

Sites, Temperature and Phenological data

Datasets were obtained from Commonwealth Scientific and Industrial Research Organisation (CSIRO), Cotton Seed Distributors (CSD) and the United States Department of Agriculture (USDA) and spanned across 16 years (planted from 2002-2017) and covered 29 field locations (Table 1). For first first square analyses data was only available from CSIRO (60 observations) and therefore was limited; the decision was therefore made to firstly evaluate functions for predicting first flower. Insights gained from these analyses were then used to assess prediction of first square. First square and flower dates were recorded when >50% of an observed group of cotton plants had a square present (subtending leaf unfurled), or flower open at the first fruiting position on the first fruiting branch of each plant. The first square (FS) and flower (FF) days after planting (DAP) were calculated as date of first square or flower minus date of planting (DAP).

For observations of first flower CSIRO and USDA data was designated dataset 1 and used for function development purposes. CSD's data (planting years 2015, 2016 and 2017) was designated as dataset 2 and was utilised solely to validate and compare the models developed. The CSD Data was chosen as the validation dataset as it spanned a greater range of regions within the industry. Dataset 1 for development included 159 data points whilst the validation dataset (2) contained 156 observations.

Temperature data (resolution to daily) was obtained from weather stations based at the site, or in the case of most CSD sites, the nearest weather station accessed via SILO patched point datasets (https://www.longpaddock.qld.gov.au/silo/). Daily mean temperature was calculated by:

$$T_{mean} = \frac{T_{min}+T_{max}}{2}$$

Table 1 Location of crops used in this study, cultivars at each year/site, and source agency for each dataset. Note that observation of first square was only recorded by CSIRO at Narrabri (60 observations).

Source	Planting Year	Location	Cultivar(s)	Observations
Development of initial rate development function				
CSIRO	Various	Phytotron, Canberra, ACT	S324, L22, V3i	261
Development of day degree targets and crop insulation factor for rate development function				
CSIRO	2002	Narrabri, NSW	189, 289B	6
	2003		189rr, 289BR	6
	2004		189rr, 289BR	6
	2007		71BR, 71BRF 70BRF, F-1B	10
	2008		71BR, 70BRF, F-1B	12
	2010		71BRF	3
	2011		71BRF	3
	2012		71BRF	3
	2013		74BRF	2
	2014		74BRF	3
	2015		74BRF	3
	2016		746BRF	3 (Total 60)
USDA	2011	Lubbock, TX (USA)	FM7180	24
	2012			28
	2014			24
	2015			8 (Total 84)
Validation				
CSD	2014	Various x 26 (NSW and QLD)	71BRF, 74BRF, 75BRF	41
	2015		714BRF, 746B3F, 748B3F, 754BRF	59
	2016		707BRF, 714BRF, 746B3F, 748B3F, 754BRF	55 (Total 155)

Model Development – First Flower

Four different modelling approaches were tested to assess prediction of flowering: (1) Current industry base temperature (12°C) threshold approach utilising a fixed target of 777 day degrees with **no** adjustment for cold shock; (2) Current industry base temperature (12°C) threshold approach utilising a fixed target of 777 day degrees with adjustment for cold shock; (3) An approach utilising both upper and base temperature thresholds (a range of thresholds were investigated with new day degree targets); and (4) A physiology based model developed explicitly from controlled temperature response experiments relying on daily average temperature providing a unit of progress towards first flower. For approaches 1, 2 and 3 a cumulative sum to a predetermined day degree target determines the prediction. In total twelve different models were tested.

(Approach 1) Current industry approach - day degree (DD) threshold 777 day degrees

$DD = \frac{(Tmax-12)+(Tmin-12)}{2}\ where\ Tmin > 12°C$ and

$DD = \frac{(Tmax-12)}{2}\ where\ Tmin < 12°C$

Where *Tmin* is the minimum temperature and *Tmax* is the maximum temperature for a day.

(Approach 2) Current industry approach adjusted for cold shock - day degree (DD) threshold 777 day degrees

$DD = \frac{(Tmax-12)+(Tmin-12)}{2}$ $\ where\ Tmin > 12°C$ and

$DD = \frac{(Tmax-12)}{2}\ where\ Tmin < 12°C$

$Day\ Degree\ Target\ Adjustment = +5.2\ \ when\ Tmin < 11\ °C$

For both approaches 1 and 2 we also generated a new day degree targets for these approaches using Dataset 1.

(Approach 3) Day Degree Estimates based on both a base and upper threshold (variable)

$DD = \frac{(Tmax\ - Base\ threshold)+(Tmin - Base\ threshold)}{2}$

$where\ Tmax > upper\ threshold,\ \ Tmax = Upper\ threeshold$

$$where\ Tmin < base\ threshold, \qquad (Tmin - Base\ threshold) = 0$$

The use of base and upper thresholds is commonly used in the USA with 15°C being used as a base temperature. There are no consistent upper thresholds used (ranging from 32 to 36°C). Interesting the derivation of the 15°C as a base temperature is more likely traced back to initial research conducted in the Canberra phytotron in the 70's by (Moraghan et al. 1968). A range of base and upper temperature threshold temperature combinations were tested (Table 2). For each combination new day degree targets were produced using Dataset 1 (CSIRO and USDA). The choice of the 17°C base and 30°C upper temperature threshold was a combination used recently in a publication by Viator et al. (2005) for boll period.

Table 2 Base and upper temperature threshold combinations tested in this study.

Name	Base	Upper
3212	12	32
3412	12	34
3612	12	36
3215	15	32
3415	15	34
3615	15	36
3017	17	30

(Approach 4) Physiology approach– applied to controlled environment studies

In this approach we related the average daily rate of progress towards flowering to the average temperature during the period of development (planting to first flower). Where the average rate of development is calculated by:

$$Average\ rate\ of\ development = \frac{1}{First\ Flower\ Days\ After\ Sowing}$$

Therefore:

$Daily\ proportion\ of\ progress\ towards\ flowering = f(daily\ Tmean)$

This response is represented by the curvilinear function:

$$Rate\ of\ development\ towards\ flowering = a(Tmean - Tmin)(Tmean - Tmax)^{v}$$

$where\ Tmean \geq Tmin,\ \ if\ Tmean < Tmin\ then\ proportion\ of\ progress = 0$

and

$$v = \frac{Tmax - Topt}{Topt - Tmin}$$

and

$$a = \frac{y_{opt}}{(Topt - Tmin)(Tmax - Topt)}$$

Topt is the maximum temperature at the asymptote of the curve, while *Yopt* is the rate at this temperature.

The reason for the choice of this function type was that this was the best fit of a regression that represented the response of first square to average temperature from planting developed from data collected in the Phytotron (Figure 1). Observations were collected for average temperatures for the period from 16 to 34 °C (data for temperature effects from the Phytotron for flowering were limited across these temperatures). The assumption was made that the response to temperature for first square would be similar for first flower because: i) First square would capture temperature effects on floral initiation that would persist through to flowering, and ii) There was a consistent rate of development for the time between square periods (period from square to flower) across temperatures (Figure 2), and this would generally not change the response. Effects were only observed at the very extremes and these were small. Again this response was measured in the Canberra Phytotron.

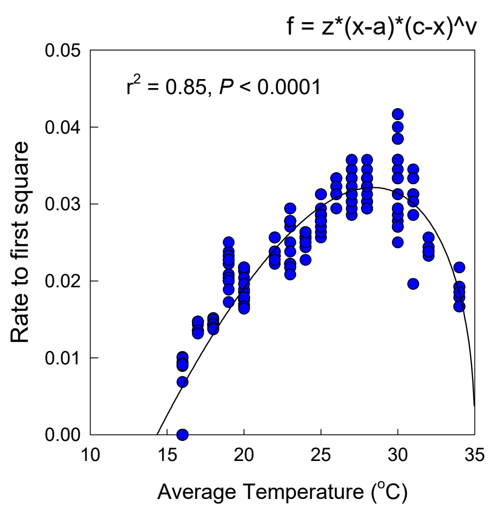

Figure 1 The shape of the function used in Approach 4 based on the temperature response of the rate to first square development generated from observations collected in the Canberra Phytotron.

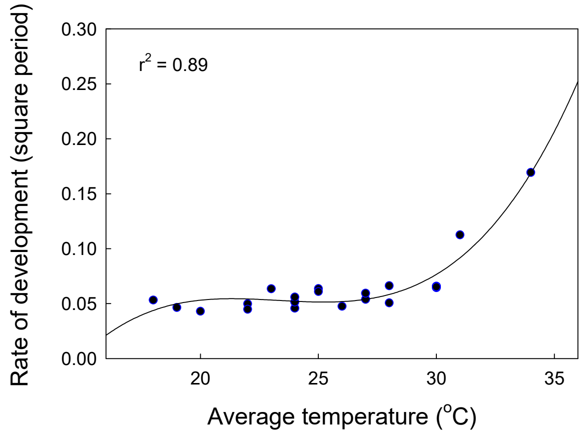

Figure 2 The response of development between first square and first flower (square period) to average temperature during this period. Observations collected in the Canberra Phytotron.

To estimate when the flowering date occurs the daily proportion of progress is summed until it reaches the value of 1. In this assessment we found that there was a bias for these results being recorded in the Canberra Phytotron compared to the field. We used Dataset 1 to create an adjustment for this bias which corrected the Phytotron data and the model was refitted adjusted data accordingly.

To generate this bias we fitted linear regressions of development rate against average temperature to both Dataset 1 and the Phytotron data separately only over the same temperature range that was measured in both datasets. We found that the slopes were not significantly different however, the intercept was (Figure 3). On average across this temperature range the field data was later by 5.234 days.

It is conceivable that temperature experienced by the crop are less than that measured in the air above the crop (or in weather stations) as the crop canopy creates a cooler environment with evapotranspiration. Essentially there may be a 'crop insulation' effect in crop canopies compared to plants grown in pots in glasshouses or controlled environments.

Potentially another reason for the bias is that the cultivars used in the Phytotron are different from those used today; therefore we tested this. Using a similar approach in calculating the bias between the Phytotron data with Dataset 1 we compared this data with older field data collected by both Constable, Bange, and Yeates (all before the year 2000). Figure 4 shows the comparison of the linear regressions of development rate against average temperature to both the old cultivar and Dataset 1 over the same temperature range that was measured in both datasets. A statistical analysis comparing the regressions showed that there were no significant differences between the responses suggesting at least with this data there were no cultivar differences.

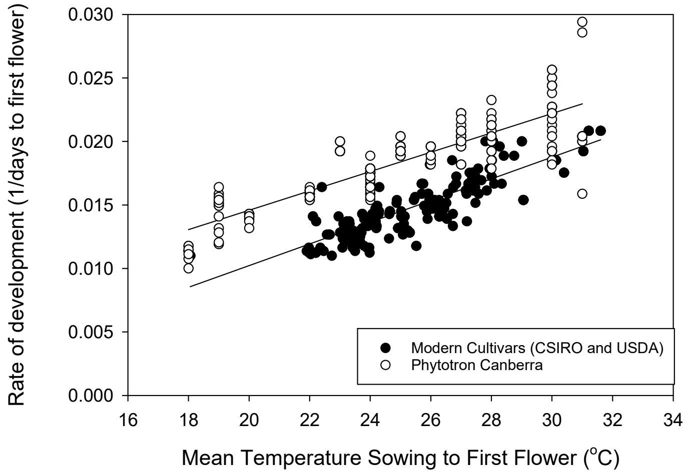

Figure 3 Regressions responses used to develop bias between Phytotron data and modern cultivars measured in the CSIRO and USDA studies. There were no significant differences in the slope of these responses.

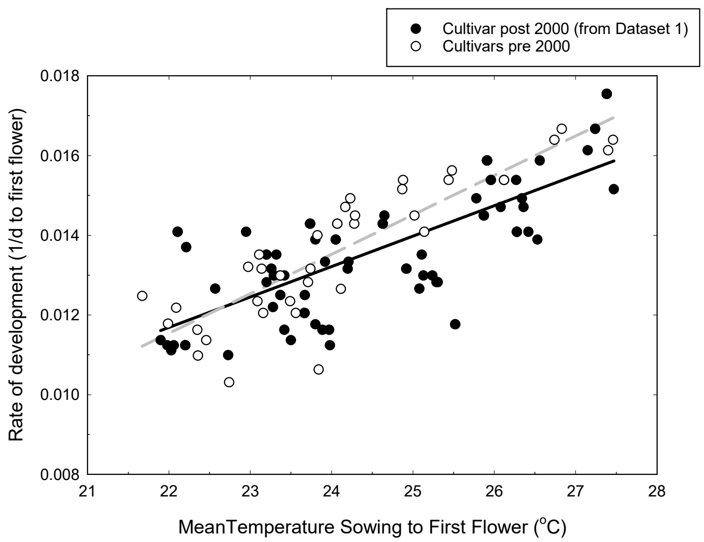

Figure 4 Regression used to test if there were differences between old (pre 2000) cultivars and those used in Dataset 1 utilised in this study. There was no significant differences in the response to temperature between the two groups of cultivars.

Figure 5 shows that results of the fitted function applied to the phytotron data (Figure 5a) (r^2 =0.74; P<0.001) and the results of the function overlying the modern field data (Dataset 1 (CSIRO and USDA)) before (Figure 5b), and with the applied bias (Figure 5c). The bias was -5.324 °C applied to average temperature of the phytotron data. Note that the temperature of the intercept denoting the average temperature at which no development occurs is 14.6°C.

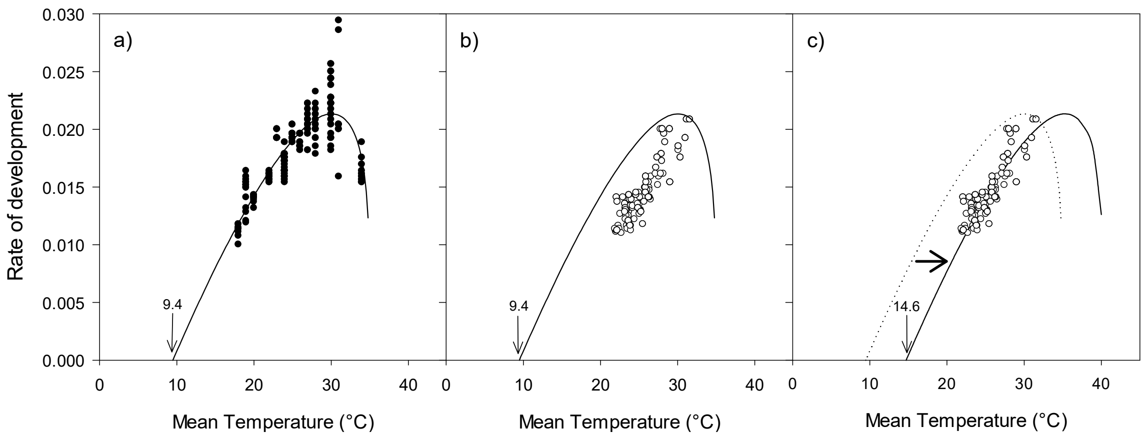

Figure 5 Curvilinear function fitted to phytotron data (a) and the horizontal shift applied to the function to account for crop insulation or modern genotype differences generated from the development dataset 1 (b and c).

The resulting response function with the bias applied is as follows:

This response is represented by the curvilinear function:

$$Rate\ of\ development\ towards\ flowering\ (d^{-1}) = 7.071 \times 10^{-4}(Tmean - 14.714)(40.234 - Tmean)^{0.24}$$

Results and Discussion

First Flower Validation

In total 12 functions were validated for their performance to predict the time of first flower (50% plants in a crop) across an independent dataset. Data used to generate functions and targets (Tables 1 and 3) represented modern cultivars used in Australia and the United States that was collected across an average temperature range from 18.05 to 31.60 °C for the period from planting to first flower. Validation of functions used data collected from Cotton Seed Distributors over three seasons in 26 different locations from Griffith to Emerald (Table 1). The temperature range experienced from planting to first flower for these crops was from 20.74 to 29.16 °C. Figure 6 is an example of the spread of the range of temperatures experienced.

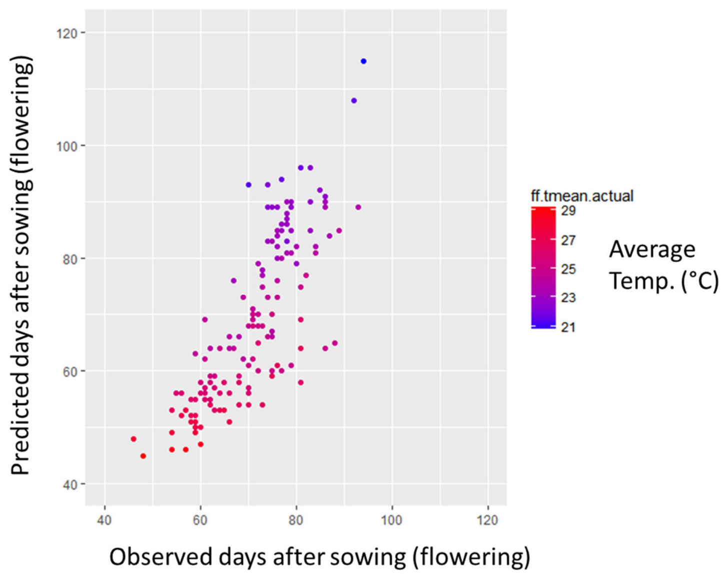

Figure 6 A graph showing the outcome of one of the models predictions highlighting the spread of validation data (Dataset 2) across the temperatures experienced from planting to first flower.

Utilising a combination of assessments described above in the methods this study was able to identify functions that performed better than the current approach used in the Australian production environment (Standard - with cold shock (Base 12), Function 2) (Table 3). In the first instance there were a considerable number of functions that had similar or better performance in terms of RMSD and MAE than Function 2 (Functions 5,6,7,8,9,10, and 12). In assessing bias (slope of the predicted versus actual data) relating to the ability of the functions to predict accurately across the validation dataset there were a considerable number of models that were better than Function 2 (Functions 3,4,5,6,7,8,9, and 11) (Table 3, Figure 7). The final assessment was to use the 'W statistic' to provide evidence that the function is able to predict accurately across a range of temperatures. The only functions that were able to satisfy the 'W statistic' test were Functions 3,4,5, and 9 (Table 3).

Overall the best functions that performed well in all assessments were Functions 5 and 9 (Figures 7 c and j respectively). Function 5 was the function developed from the phytotron Data (with adjustment) while Function 9 was the function that had a base temperature of 15 °C and an optimum temperature of 32 °C. A useful outcome of the phytotron analysis was that the intercept at which there was no cotton development towards flowering was 14.6 °C supporting the use of 15 as a base temperature in the other successful function (Function 9). Similarly when investigating where the slope of Figure 3c the rate of development also slows around 32 °C.

Importantly when assessing the current function with new targets (Functions 3 and 4) there was no improvement in the predictability (Table 3, Figure 7 d and e), suggesting that the base temperature of 12 °C is not appropriate and an optimum temperature or slowing in development at high temperatures is needed. These new functions have also been developed over a greater range of genotypes compared to that of the original function developed by Greg Constable; this adds additional robustness to the outcomes generated.

Other benefits of the new functions identified are that they do not require the use of a cold shock adjustment to enable the prediction of outcomes. This builds on earlier work by Bange and Milroy (2004) where they could not find strong evidence of a cold injury effect delaying cotton development.

Table 3 RMSD, MAE, slope, and W statistic for the functions evaluated in this study. Also presented are the targets used by the functions to predict outcomes that were generated using Dataset 1. Note that the highlighted tests (in green) of models are the best functions ones performing in that category.

Function name and number	RMSD	MAE	Slope	W-test	W - prob	Function DD target
1. Standard - no cold shock (Base 12)	10.84	9.20	0.63	0.98	0.02	777
2. Standard - with cold shock (Base 12)	6.79	5.21	0.63	0.98	0.06	777
3. Standard (current data) - no cold shock (Base 12)	9.13	7.41	1.20	0.99	0.45	867
4. Standard (current data) - with cold shock (Base 12)	9.79	7.56	1.20	0.99	0.44	867
5. Phytotron adjusted rate approach	7.99	6.60	0.85	0.99	0.12	1
6. Variable - Base 12, Upper 32	6.47	4.82	0.74	0.98	0.01	860
7. Variable - Base 12, Upper 34	6.70	4.99	0.72	0.98	0.01	860
8. Variable - Base 12, Upper 36	6.70	4.99	0.72	0.98	0.01	860
9. Variable - Base 15, Upper 32	7.30	5.78	0.92	0.99	0.28	656
10. Variable - Base 15, Upper 34	6.47	5.30	0.63	0.98	0.04	648
11. Variable - Base 15, Upper 36	8.04	6.17	0.95	0.98	0.02	656
12. Variable - Base 17, Upper 30	6.12	4.69	0.57	0.98	0.02	481

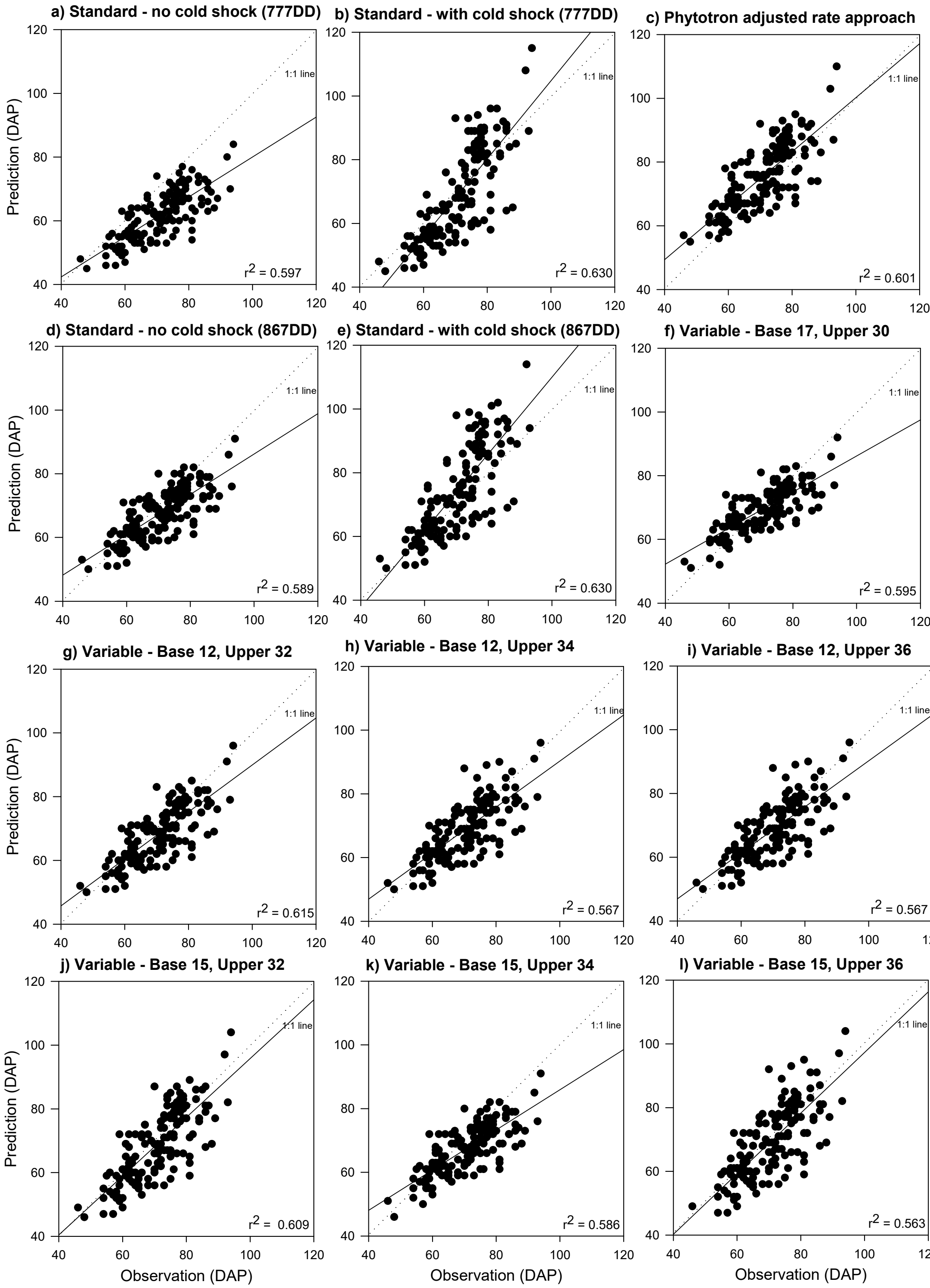

Figure 7 Predicted versus Observed dates for first flower using the validation datasets (CSD's 2014, 2015 and 2016 plantings – Dataset 2). Each graph shows a linear regression fitted along with the 1:1 shown to demonstrate predictive capability. DAP is days after planting.

First Square Validation

Following the insights gained from the analysis of functions using the broader datasets associated with first flower, an analysis of specific functions types generated for first flower were applied to predict first square (50% of plants). The reason for choosing a similar function type is that for ease of application in industry, as a similar function type will avoid confusion when applied across phenological stages. As stated previously first square measurements were only taken in the Narrabri component of Dataset 1. This meant that here we are only able to demonstrate the performance of the functions on a limited dataset, and the assessments are therefore not independent as the same data was used to generate targets as for assessing the functions.

From the analyses above four function types were assessed:

1. The standard Australian function (Function 2, Table 2, Approach 2) using a target day degree value of 551 DD;
2. A variable function with a base temperature of 15 °C and an optimum of 32°C (Function 10, Table 2, Approach 3). The day degree target generated using the field observations was 452 DD.
3. The physiology rate response (Function 5, Table 2, Approach 4). Using the field observations the average proportion of time when first square appeared between the period between planting first flower was 0.63. This was used as the target value.

Plotting the predicted versus observed days after planting for first square showed that the slope of the regressions were closer to 1 for the variable and physiology rate function compared to the standard function (Figure 8). In addition the RMSD and MAE were also better for these functions compared to the standard function.

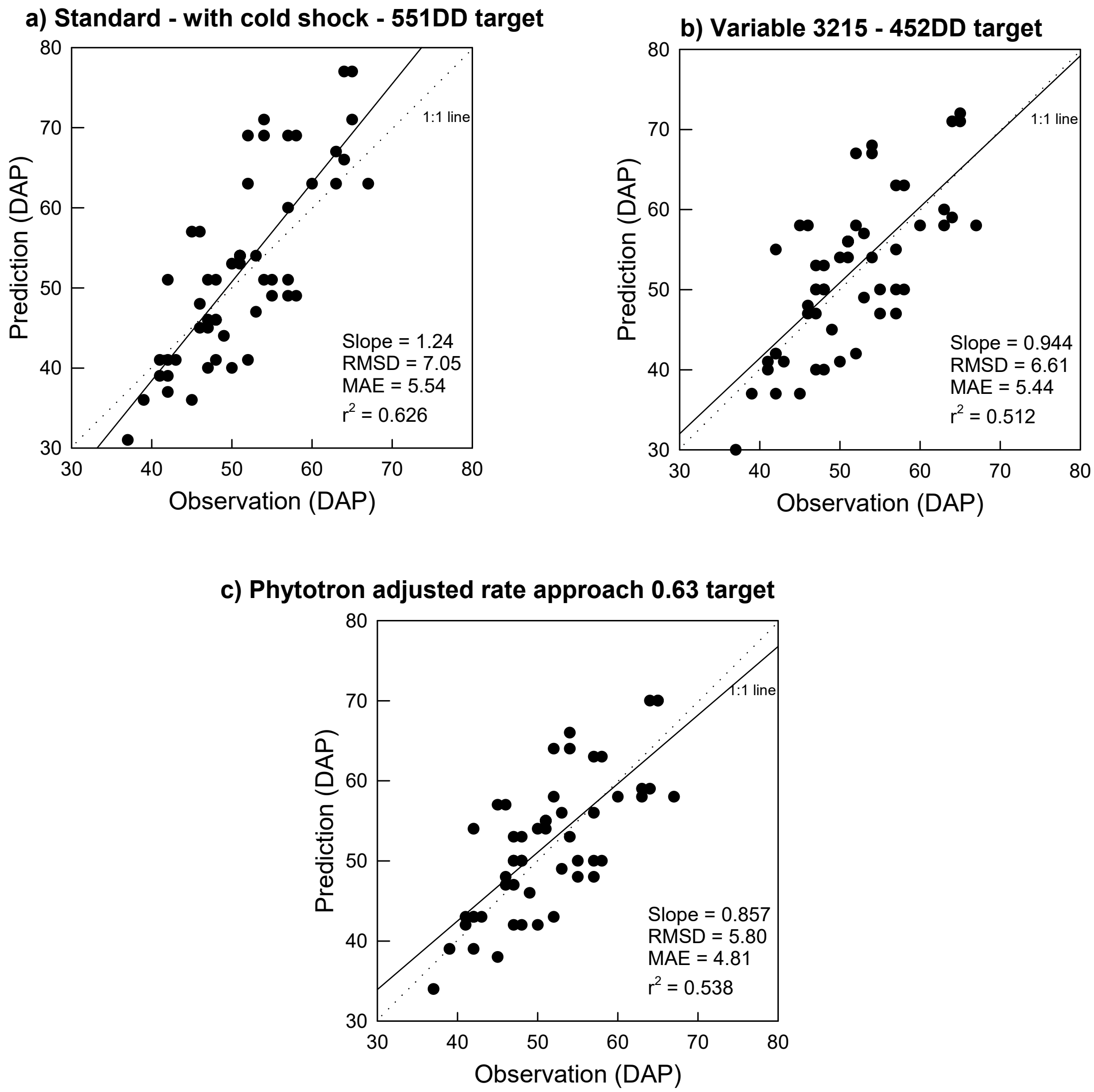

Figure 8 Predicted versus Observed dates for first square using the CSIRO dataset. Each graph shows a linear regression fitted along with the 1:1 shown to demonstrate predictive capability. DAP is days after planting. RMSD and MAE are also presented to compare functions.

Conclusions

This study was able to demonstrate that there were improvements in the predictability of time of first square and first flower measured in crops. Two functions were able to better predict these phenological stages compared to the existing function used currently in the Australian industry (Constable and Shaw, 1988). The best performing functions were a variable temperature day degree function that used a base temperature of 15 °C and an optimum of 32 °C, and a physiological rate function that reflected similar temperature characteristics as the variable function. The use of these functions should be considered in the development of new cotton crop predictive capabilities as they will be able to account for more temperature extremes (high and low, that maybe more prevalent in a changing climate) and where cotton production moves into new regions.

The analysis of functions here also support the use of a base temperature of 15 °C (60 °F) used in USA cotton systems. Ironically it appears that the USA cotton systems have used this temperature based on early work undertaken in Australia in the 1960's by Moraghan et al. (1968). The best functions derived here will also be simpler to apply as no adjustments for cold shock/injury need to be applied to adjust day degree targets.

Future research could look to refine predictive capabilities by accessing hourly temperature data which is becoming more accessible as more growers employ the use of automatic weather stations. Another opportunity also exists in exploiting the use of continuous thermal measures of crop temperature (e.g. Canopy temperature sensors used in irrigation) to more precisely measure the crop in a specific field. Recent assessments of the relationships of canopy temperature to air temperature highlight that there can be substantial differences between the two; this highlights a source of error associated with using current day degree approaches. The use of plant based temperature sensing would also open up the opportunity to predict development of crops that are stressed. Current day degree models assume no stress on crops.

References

Bange, M.P. and Milroy, S.P. (2004). Impact of short-term exposure to cold night temperatures on early development of cotton (*Gossypium hirsutum* L.). Australian Journal of Agricultural Research 55:655-644.

Constable G.A. and Shaw A.J. (1988). Temperature requirements for cotton. Agfact P5.3.5. (Division of Plant Industries, New South Wales Department of Agriculture).

Miranda, C. Santesteban, L.G. and Royo, J.B. (2013). Evaluation and fitting of models for determining peach phenological stages at a regional scale. Agricultural and Forest Meteorology. 178-179:129-139.

Moraghan, B.J., Hesketh, J. and Low, A. (1968). Effects of temperature and photoperiod on floral initiation among strains of cotton. Cotton Growing Review. 45:91-100.

Viator, R. P., et al. (2005). "Predicting cotton boll maturation period using degree days and other climatic factors. Agronomy Journal 97(2): 494-499.

CONTACT US

t 1300 363 400
+61 3 9545 2176
e csiroenquiries@csiro.au
w www.csiro.au

AT CSIRO, WE DO THE EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today – for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australia's catalyst for innovation.

CSIRO. WE IMAGINE. WE COLLABORATE. WE INNOVATE.

FOR FURTHER INFORMATION

CSIRO Agriculture and Food
Michael Bange
t +61 2 6799 1540
e first.last@csiro.au
w www.csiro.au/Agriculture