

Development and evaluation of a field-based high-throughput phenotyping platform

Pedro Andrade-Sanchez^{A,E}, Michael A. Gore^{B,C}, John T. Heun^A, Kelly R. Thorp^B,
A. Elizabete Carmo-Silva^{B,D}, Andrew N. French^B, Michael E. Salvucci^B and Jeffrey W. White^B

^ADepartment of Agricultural and Biosystems Engineering, University of Arizona, Maricopa Agricultural Center, 37860 W. Smith-Enke Road, Maricopa, AZ 85138, USA.

^BUS Department of Agriculture, Agricultural Research Service, Arid-Land Agricultural Research Center, 21881 North Cardon Lane, Maricopa, AZ 85138, USA.

^CPresent address: Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY 14853, USA.

^DPresent address: Rothamsted Research, Plant Biology and Crop Science Department, Harpenden, Hertfordshire, AL5 2JQ, UK.

^ECorresponding author. Email: pandrade@ag.arizona.edu

Abstract. Physiological and developmental traits that vary over time are difficult to phenotype under relevant growing conditions. In this light, we developed a novel system for phenotyping dynamic traits in the field. System performance was evaluated on 25 Pima cotton (*Gossypium barbadense* L.) cultivars grown in 2011 at Maricopa, Arizona. Field-grown plants were irrigated under well watered and water-limited conditions, with measurements taken at different times on 3 days in July and August. The system carried four sets of sensors to measure canopy height, reflectance and temperature simultaneously on four adjacent rows, enabling the collection of phenotypic data at a rate of 0.84 ha h⁻¹. Measurements of canopy height, normalised difference vegetation index and temperature all showed large differences among cultivars and expected interactions of cultivars with water regime and time of day. Broad-sense heritabilities (H^2) were highest for canopy height ($H^2 = 0.86–0.96$), followed by the more environmentally sensitive normalised difference vegetation index ($H^2 = 0.28–0.90$) and temperature ($H^2 = 0.01–0.90$) traits. We also found a strong agreement ($r^2 = 0.35–0.82$) between values obtained by the system, and values from aerial imagery and manual phenotyping approaches. Taken together, these results confirmed the ability of the phenotyping system to measure multiple traits rapidly and accurately.

Additional keywords: cotton, genetics, *Gossypium barbadense*, phenomics, proximal sensing.

Received 4 May 2013, accepted 18 July 2013, published online 5 September 2013

Introduction

In order for crop improvement to satisfy the expected demand for increased yield potential and abiotic stress tolerance, the plant science community must connect genotype to phenotype with unprecedented efficiency. Perhaps the greatest challenge for plant breeding and genetics research in the 21st century is establishing the capacity for rapidly and accurately phenotyping very large numbers of field-grown plants (Montes *et al.* 2007; Houle *et al.* 2010; Furbank and Tester 2011). To date, the development of high-throughput phenotyping systems for plants has largely focussed on measuring traits of individual plants in greenhouses or growth chambers. However, many phenotypic responses of interest for crop improvement, especially those related to yield potential and abiotic stress tolerance, involve suites of traits that are best measured as expressed among communities of plants that grow in agronomically relevant edaphic and climatic conditions. Furthermore, field-based systems are more readily incorporated into applied plant breeding programs. Thus, there is growing

interest in adapting agricultural machinery and electronic sensors for field-based high-throughput phenotyping (Montes *et al.* 2007; White *et al.* 2012). Potential applications are mainly envisaged for genetic research (e.g. in detection of quantitative trait loci, QTL) and crop improvement but also include monitoring of the crop response to soil and management variability (i.e. precision agriculture).

Various vehicle-based high-throughput systems have been used or proposed for phenotyping plants in the field. Ruixiu *et al.* (1989) used a three-wheeled cart to position multiple ultrasonic proximity sensors around a single row of a crop. Rundquist *et al.* (2004) described a tracked vehicle with a boom extendable to 12 m. Montes *et al.* (2007) proposed collecting spectral data either from tractor- or harvester-mounted reflectance sensors, and McCarthy *et al.* (2010) described a machine vision system that could measure internode length in cotton (*Gossypium* spp.) plants at a speed of 0.2 m s⁻¹ along a row. These examples demonstrate deployment of important single sensor types, but

they are substantially limited when attempting to assess multiple-traits, which is needed for crop improvement.

The system of Lan *et al.* (2009) integrated sensors for crop canopy height, leaf area index (LAI), normalised difference vegetation index (NDVI), multispectral imaging and hyperspectral reflectance, but the system recorded data for only one crop row per pass. Scotford and Miller (2004) tested a passive radiometer and one ultrasonic sensor mounted in a laterally deployed frame attached to the rear three-point hitch of an agricultural tractor, showing that the combined data resulted in improved estimation of LAI. Comar *et al.* (2012) implemented a rear-mounted platform with a variety of instruments including a global positioning system (GPS) antenna, a passive spectrometer, an irradiance probe and a digital camera. The handcart described by White and Conley (2013) was capable of positioning multiple sensors over two crop rows but had clearance limitations and would be restricted by the labour required to move the cart on a continuous basis, especially under high field temperatures.

Although most of these systems had elements that would improve the acquisition of phenotypic data, none appeared capable of providing the needed throughput for the multiple data types that are essential for efficient field-based plant phenotyping. A further concern is that the evaluations, if given, seldom include estimates of throughput or of the heritability of the measured traits, both of which are key concerns for adoption in plant genetics research or crop improvement. Fortunately, current technologies provide an unprecedented array of hardware and software solutions that appear capable of providing the requisite high throughput. Examples for instruments largely relate to the increasing power of electronic components through greater integration of functions and reduction in size, with concomitant reductions in cost and power consumption. Advances in vehicle function include real-time control mechanisms, such as GPS-enabled automatic tractor or implement steering, and controllers for vertical boom position adjustments.

Our approach to sensor deployment falls in the category of proximal sensing, contrasting with the remote deployment of sensors using aerial or satellite platforms (Fussell *et al.* 1986). In this context, a plant phenotyping system should allow for frequent deployment of replicated sets of different sensors at close proximity to plants, enabling the simultaneous phenotyping of several traits on plants from multiple adjacent rows. Enabling a plant phenotyping system to record multiple types of data in a single pass would not only increase throughput, but would also enable more accurate and comprehensive specification of phenotypes (Scotford and Miller 2004; White *et al.* 2012). In addition, the system should adhere to requirements for sensor deployment, as outlined in Milton's landmark review of field spectroscopy (Milton 1987), including the following criteria: (1) sensors should remain at a fixed geometry to assure repeatability; (2) sensors should be deployed at least 1 m above the top of the plant canopy; and (3) sensors should view plants close to the solar principal plane (i.e. parallel with the azimuthal direction of the sun) regardless of platform orientation.

Herein, we describe and evaluate a system that possesses many of the functionalities needed for efficient phenotyping of plants in the field. The evaluations involved measuring morphological and physiological responses of Pima cotton (*Gossypium*

barbadense L.) to well watered and water-limited conditions, and by comparisons of these measurements with aerial imagery and manual phenotyping approaches. The specific traits measured were canopy height, reflectance and temperature. Although individual traits such as height have interest for breeding (e.g. in relation to harvestability and harvest index), we emphasise that realising the full benefit of a platform capable of simultaneously measuring multiple canopy traits will probably require integrated analysis of the data streams to describe more complex crop traits such as growth rates and transpiration efficiency (White *et al.* 2012).

Materials and methods

Experimental design

The field experiment was conducted in the summer of 2011 at the Maricopa Agricultural Center of the University of Arizona (33°04'N, 111°58'W, elevation 360 m) in Maricopa, Arizona. The location is within an irrigated production area in the low desert that experiences less than 100 mm of rainfall during the April to September growing season. Meteorological data for the experimental field site were obtained from the Arizona Meteorological Network (ag.arizona.edu/azmet/index.html; Brown 1989) with an automated weather station 270 m from the field. The soil was a Casa Grande sandy loam (fine-loamy, mixed, superactive, hyperthermic Typic Natrargids).

We evaluated a set of 25 Pima cotton (*Gossypium barbadense* L.) cultivars that captured a wide range of genetic variability for heat and drought tolerance. The set included Pima cotton cultivars released over a period of 90 years (1918–2009) by breeding programs in Arizona (Old Pima, Amsak, Pima 32, Pima S-1, Pima S-2, Pima S-3, Pima S-4, Pima S-5, Pima S-6, P53, P62, Pima S-7, P70, P73, P76, 8810, 89590, 93260, 94217, 94220, PSI113 and PSI425), heat-sensitive Sea Island type cultivars from the Caribbean (Sea Island St. Vincent V-135 and Monseratt Sea Island) and a recently released commercial Pima cotton variety (Deltapine 360). For the set of 25 cultivars, a well watered (WW) and water-limited (WL) trial was arranged as a $5 \times 5 \alpha(0,1)$ lattice design with four replications for a total of 200 plots. Experimental units were one-row plots 8.8 m long with a 1.02-m inter-row spacing. There was a 0.61 m alley at the end of each plot. Plots were thinned to ~ 4.1 plants m^{-2} . To reduce border effects, one-row plots of the commercial Pima cotton variety PhytoGen 800 (PHY 800) were planted on all sides of each 5×5 replicate. With the inclusion of border plots, the total field consisted of 24 rows that had an inter-row spacing of 1.02 m and a length of 109.73 m (Figure S1, available as Supplementary Material to this paper).

Furrow irrigation was used to germinate the seed and establish the crop. Subsurface drip irrigation was installed before planting, and plots were drip-irrigated after crop establishment. A daily soil water balance model calculated for the cotton root zone was used to schedule the timing and amount of subsurface drip irrigation, following the Food and Agriculture Organisation-56 approach (Allen *et al.* 1998). Soil water balance inputs included estimated daily evapotranspiration determined from the FAO-56 dual crop coefficient procedures, metered irrigation depths and meteorological data obtained from the Arizona Meteorological Network weather station on the research farm. Soil water

depletion was also monitored through weekly soil moisture measurements to a depth of 1.5 m. These measurements were used to adjust the modelled soil water balance.

Prior to 8 July, both WW and WL plots received equal irrigation amounts. Specifically, when plots were at ~35% soil water depletion, subsurface irrigations was applied to refill the water content of the cotton root zone to field capacity. To minimise the interaction of flowering time and drought stress, we imposed the first WL treatment when more than half of the plots had reached the first flower stage (i.e. 50% of the plants within a plot had at least one visible flower). Starting on 8 July and continuing until 6 September, the WL plots received one-half of the irrigation amounts applied to the WW plots. The final irrigation was on 6 September, 17 days before the cotton plants were defoliated.

Plant phenotyping system

The plant phenotyping system was built on a LeeAgra 3434 DL open rider sprayer (LeeAgra, Lubbock, TX, US). The vehicle had a maximum height clearance of 1.93 m, thus providing minimal disturbance to the plants. A boom was attached to the front end of the tractor frame to provide mechanical support for sensors, data loggers and other instrumentation components such as enclosure boxes and the cables needed for plant phenotyping. The boom was supported by two lift arms that allowed vertical movements via rephasing-type hydraulic cylinders. Parallel bars in the lift arms maintained the boom on a horizontal plane as it moved vertically. Boom stability was enhanced by the wide separation (2.31 m) between each pair of lift arms. The high clearance vehicle had an isolated secondary power source to supply the electronic components with a 12-V direct current power for plant phenotyping measurements. The four sets of sensors and accompanying electronic instrumentation had an average current consumption of 2.38 A.

We deployed three types of sensors (described below) for measuring plant canopy height, temperature and reflectance. The sensors were arrayed in four sets on the front boom, which allowed phenotypic measurements to be collected simultaneously from four adjacent rows of cotton plants (Fig. 1). The forward position of the instrumentation boom allowed measurements with no disruption to the plants caused by the vehicle platform. Sensor brackets were horizontally and vertically adjustable, thus allowing the sensors to be positioned above the crop canopy at the centre of a row and at the desired height. Sensor positions, along with direction of travel, were considered when determining the geographical coordinates of each collected data point.

Plant canopy height (in mm) was measured with a sonar proximity sensor that used a short-range (125–3000 mm) Pulsar dB3 transducer (Pulsar Process Measurement Ltd, Malvern, UK) with a resonant frequency of 125 kHz, a 19-mm diameter radiating face and a 10° beam angle. A Pulsar Black Box Level 130 controlled the dB3 transducer, applied signal processing and provided a 0–20 mA output proportional to distance. The current output signal was converted to a proportional voltage through the use of a 100- Ω metal foil high-precision resistor. An empirical calibration procedure was used to convert the output signal to units of distance.

Sonar and GPS antenna elevation data were combined to calculate a measure of canopy height (*CH*). Elevation, or altitude, is expressed as meters above mean sea level and was recorded from the GGA string output by the GPS. This string is one of multiple message formats defined by the National Marine Electronics Association (NMEA, Severna Park, MD, USA) in the serial communications protocol NMEA 0183. The first step was conducting an elevation survey of field beds immediately after planting to determine a baseline reference elevation (E_{ref}) for each plot. Canopy height (mm) was calculated after processing using the established reference elevations, the sonar transducer distance output (d , in mm) and the elevation of the sonar transducer face (E_{sonar}). E_{sonar} was defined by subtracting the distance between the GPS antenna and the sonar sensor from the GPS antenna elevation. The position of the sonar sensor with respect to the GPS antenna was manually set and recorded before system deployment in the field. Canopy height calculations were based on the following equation:

$$CH = E_{\text{sonar}} - E_{\text{ref}} - d. \quad (1)$$

Canopy temperature ($^\circ\text{C}$) was measured with two Apogee SI-121 infrared radiometer (IRT) sensors (Apogee Instruments, Logan, UT, US) that had a narrow field of view (18° half-angle). The sensor output was converted to temperature values after applying polynomial calibration equations supplied by the manufacturer. IRT sensors were installed in pairs to improve their discrimination of plant canopy temperatures from bare soil temperatures. For each pair of IRT sensors, one sensor was installed in the vertical direction pointing downward (nadir), and the other sensor was mounted at 30° from the vertical axis in a forward-facing (down the row) direction. With field calibration, this dual-view angle configuration can be used to transform composite IRT temperature data into distinct vegetation and soil temperatures (Kimes 1981). For 2011, this calibration was not done, meaning that only the nadir viewing IRT sensors were analysed. This meant that additional screening of the temperature data using statistical methods had to be performed to remove outlier observations that resulted from passing over large gaps in vegetation.

We measured the plant canopy reflectance (ρ) of each plot with a Crop Circle ACS-470 multi-spectral crop canopy sensor (Holland Scientific, Lincoln, NE, US). The sensor had three channels for light measurements over visible and near infrared (NIR) wavelengths. In this study, we chose filters with centre wavelengths of 670 nm, 720 nm and 820 nm that corresponded to red, red-edge and NIR light, respectively. Red and red-edge filters had a bandwidth of 10 nm, and that of the NIR filter was ~60 nm. Spectral data were recorded with a Holland Scientific GeoScout GLS-420 datalogger. Canopy reflectance data in the NIR (ρ_{NIR}) and red (ρ_{red}) regions were used to calculate normalised difference vegetation index (NDVI) as follows:

$$NDVI = \frac{(\rho_{\text{NIR}} - \rho_{\text{red}})}{(\rho_{\text{NIR}} + \rho_{\text{red}})}. \quad (2)$$

Placing sensors in the vertical direction is of prime importance to the performance of the phenotyping system. Adjustments to sensor height were made to restrict the field of view of individual sensors to avoid cross-contamination with adjacent rows. Plant



Fig. 1. (a) Front view of the plant phenotyping system during field deployment in Maricopa, Arizona. Close-up views of instruments include: (b) the sonar proximity sensor, (c) the infrared radiometer sensor, (d) the GPS-RTK receiver-antenna and (e) the multi-spectral crop canopy sensor.

height sensors were positioned 30 cm above the tallest plants, well above the minimum sensing range (blanking distance) of 12.5 cm. IRT sensors were positioned 15 cm above the plant canopy. Spectral sensors were fixed at 170 cm above the top of the planting bed.

The sensor platform of the phenotyping system moved at a constant speed of 0.75 m s^{-1} while simultaneously scanning four rows of plots. Measurements were georeferenced by simultaneously recording position and elevation data from a Global Navigation Satellite System real-time kinematics (RTK) GPS receiver (A320 Smart Antenna, Hemisphere GPS, Scottsdale, AZ, US) via the NMEA GGA string. A rover receiver was mounted at the centre of the boom and programmed to output position data at 1 Hz. Serial information from the GPS receiver was sent on a RS-232 network to three data loggers: GeoScout GLS-420, CR1000 and CR3000 (Campbell Scientific, Logan,

UT, US). A separate receiver (A321 Smart Antenna, Hemisphere GPS) was used as a base station unit to broadcast high-precision correction data.

Due to the number of differential and single-ended voltage channels required for the sensors, data generated by the IRT sensors were stored on the CR3000, whereas data generated by the ultrasonic proximity sensors were stored on the CR1000. CR-Basic (Campbell Scientific) code was written to read and record analogue sensor signals and GPS serial data at a sampling frequency of 1 Hz.

Field performance of the plant phenotyping system

Field performance of the system was evaluated using two criteria. First, field efficiency was calculated as the ratio of productive time under field conditions (phenotyping) to the total time in the field

(phenotyping plus turning and idle travel). This measure of efficiency was calculated based on timestamps extracted from GPS strings collected during deployment. Second, we calculated the ability of the system to generate data on a time and area basis. This was calculated by tabulating the size (in MB) of electronic data files, in their native format, which were generated when using the system. The system was driven over the field on 21 July (0800–0900 hours and 1300–1400 hours Mountain Standard Time (MST)), 4 August (0700–0800 hours, 1100–1200 hours and 1500–1600 hours MST) and 18 August (0700–0800 hours, 1100–1200 hours and 1500–1600 hours MST). Phenotypic data collection times varied due to instrument preparation constraints but attempted to assess early morning (low stress), midday (moderate to severe stress) and midafternoon (maximum stress) conditions with respect to plant water deficits. Weather conditions during measurements are summarised in Table 1.

Helicopter flights

For verification in this research context, comparable visible and near infrared (VNIR) and thermal infrared (TIR) data from an alternative sensing platform were collected from an externally mounted, helicopter-based (Hiller UH-12, Palo Alto, CA, USA) 8-kg instrument package containing cameras, power supplies and electrical connections, all mounted on a 305 × 381 × 6.35 mm aluminium plate. Data were collected on board with two laptop computers (T-61 IBM-Thinkpad, Lenovo, Singapore). The flights were conducted at ~600 m above the ground, yielding image data at a 0.5-m resolution. Two camera systems were used. A FLIR SC645 camera (FLIR Systems, Wilsonville, OR, US) collected 16-bit, 640 × 480 pixel thermal images at 1 Hz with a manufacturer-rated accuracy of 2°K. For the VNIR images, two parallel mounted μ Eye monochromatic complementary metal–oxide–semiconductor (CMOS) cameras (IDS Imaging Development Systems GmbH, Obersulm, Germany) were used. The cameras collected 8-bit 768 × 576 pixel images with a rolling shutter format at 10 Hz. Vegetation index data were obtained by mounting interference filters on each 14-mm focal length lens: red (670 nm, 10 nm bandwidth) on one camera and NIR (760 nm, 10 nm bandwidth) on the other. Both TIR and VNIR image datasets were georegistered to Universal Transverse Mercator coordinates using United States Geological Survey (USGS) orthophotographic maps (www.nationalmap.gov/ortho.html).

Image processing required different procedures for the two camera systems. The TIR data collected on 21 July and 18 August were processed using FLIR ExaminIR software (www.flir.com),

converted to public format, and corrected for atmospheric effects. This latter step used North American temperature–humidity profile National Centers for Environmental Prediction (NCEP) data from the National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory (www.esrl.noaa.gov), spectral response function data from FLIR Systems and the radiative transfer modelling software package MODTRAN5 (www.modtran5.com). Aircraft TIR data collected on 21 July had to be corrected for diurnal changes, since the overpass time (1058 hours MST) was significantly different from the plant phenotyping system data collection times (1300 hours MST). Aircraft TIR data collected on 18 August (1100 hours MST) were simultaneous with ground observations. Using a sinusoidal temperature model (Göttsche and Olesen 2001), TIR data collected by the plant phenotyping system were reprojected to 1058 hours MST before calibration analysis. The resulting outputs were plant canopy temperature images. The VNIR data, collected on 21 July (1058 hours MST), 4 August (1112 hours MST) and 18 August (1100 hours MST), were processed with in-house C++ software built on the Qt framework (www.qt-project.org). Custom C-language programs were used to convert VNIR digital counts to spectral reflectances and NDVI. For all three flights, four 8 m × 8 m canvas reference tarpaulins with nominal reflectances of 4%, 8%, 48% and 64% were deployed in the field for calibration.

Plant phenotyping with hand-held instruments

Plant height (mm) was measured as the distance from the soil line of the plant to the terminal bud with a calibrated bar-coded ruler. A single plant of median height was visually selected and measured in the morning (0900–1000 hours MST) on 22 July, 4 August and 19 August. To reduce the influence of plot edge effects, only plants within the internal 6.40 m of each plot were considered for measurement.

A subset of five representative cotton cultivars was selected for more detailed physiological measurements: Monseratt Sea Island, Pima S-6, P62, 89590 and PSI 425. The leaf relative water content (RWC) was determined as previously described (Carmo-Silva *et al.* 2012). Samples consisting of two 2-cm² leaf discs from a young fully expanded leaf were collected in the morning (0900–1030 hours MST) on 21 July, 4 August and 18 August. Two plants were sampled from each of the four replicate plots of each cultivar in the WW and the WL plots. Similar to plant height, only plants within the internal 6.40 m of each plot were measured for leaf RWC.

Table 1. Ambient conditions at the time of phenotypic data collection in Maricopa, Arizona, in 2011
Values are for means (temperature and humidity) or integrals (radiation) at hourly intervals (Mountain Standard Time (MST))

	July 21			August 4		August 18		
	0800– 0900 hours	1300– 1400 hours	0700– 0800 hours	1100– 1200 hours	1500– 1600 hours	0700– 0800 hours	1100– 1200 hours	1500– 1600 hours
Air temperature (°C)	30.7	37.4	25.3	35.7	39.5	29.6	38.3	42.6
Relative humidity (%)	45.8	19.8	74.7	29.5	22.0	45.2	24.7	15.7
Solar radiation (MJ m ⁻² h ⁻¹)	1.11	2.27	0.38	2.88	2.97	0.24	2.81	2.91

Geospatial data processing

The data storage and transfer limitations of the plant phenotyping system needed several aspects of data processing to be relegated to postprocessing. In particular, geospatial tools were required to locate and summarise data accurately within field plot boundaries. Because manual geoprocessing was time-consuming, labour-intensive and prone to error, geoprocessing tasks were automated with software scripts that would be readily transferable to other vehicle-based phenotyping systems. To that end, algorithms and user interfaces were developed as a plug-in for the open-source Quantum Geographic Information System (GIS) environment (www.qgis.org). The plug-in incorporated two main custom tools: the preprocessor and the geoprocessor.

The two main functions of the preprocessor were (1) to convert latitude and longitude coordinates to the Universal Transverse Mercator coordinate system, and (2) to calculate coordinate transformations from the GPS-RTK receiver to the sensor positions depending on the direction of travel. Proper transformations were important for this high-throughput phenotyping application, because the sensors at each position along the boom collected data from different treatment plots. The preprocessor tool was designed for flexibility and could accommodate different column orderings for latitude, longitude, vehicle heading and sensor values. Offsets (in m) from the GPS-RTK receiver to each sensor position were required to be inputted by the user.

The geoprocessor analysed sensor data within plot boundaries delimited by rectangular polygons, which required the availability of an accurate plot boundary map, loaded as a polygon shapefile within the GIS software. The geoprocessor calculated summary statistics, such as means and s.d., for sensor data collected within each plot boundary, and the statistics were appended to the plot boundary layer as attributes of individual polygons. In contrast to the default geospatial tools commonly provided with GIS software, the geoprocessor iteratively summarised the sensor data independently for each feature in the plot boundary layer. The geoprocessor also permitted us to assign a plot identifier to each data point, and the plot identifiers were appended to the sensor data layer. In this study, the geoprocessor was used to export raw phenotypic data points from only within the central 6.40 m of each plot to minimise border effects.

Statistical analyses

We analysed plant canopy height, NDVI and temperature data from the 25 Pima cotton cultivars. Initially, a mixed linear model was fitted separately for each trait with the MIXED procedure in SAS for Windows ver. 9.3 (SAS Institute, Cary, NC, USA). Each model included one of the traits as the dependent variable. The cultivar main effect and water regime main effect, and their two-way interaction were included in the model as fixed effects; replicates nested within water regime and blocks nested within replicate were included as random effects. Outliers were detected by examining the Studentised deleted residuals (Kutner *et al.* 2004) obtained from these fitted models. We then calculated means for each plot by averaging the trait values that had been screened for outliers. The Box-Cox procedure (Box and Cox 1964) was used in SAS ver. 9.3 (SAS Institute) to find the most

appropriate transformation of these means that corrected for non-normality of the error terms and unequal variances. Data transformation was implemented if the power parameter was not equal to 1 (i.e. $\lambda \neq 1$). The transformed means of each trait were then used as the dependent variables in all subsequent statistical analyses.

For each of the 3 days, a univariate mixed linear model was fitted with the plot means of one of the three traits (i.e. canopy height, temperature or NDVI) as the dependent variable. The main effects of cultivar, water regime and time of day, as well as their corresponding interaction terms, were fitted as fixed effects. Additionally, replicates nested within water regime and blocks nested within replicate were included as random effects. The REPEATED statement was used in PROC MIXED to model the correlation structure of repeated measurements of a trait on the same experimental plots over a period of time. Likelihood ratio tests were conducted to remove nonsignificant random terms ($\alpha=0.05$) from the final model. Degrees of freedom were calculated with the Satterthwaite approximation. Least square means were obtained with the LSMEANS statement in PROC MIXED. The parameter estimates were back-transformed to report results in the original scale.

Broad-sense heritability on an entry-mean basis (H^2) or repeatability (Piepho and Möhring *et al.* 2007) was estimated with a mixed linear model fitted separately for each trait with PROC MIXED in SAS ver. 9.3 (SAS Institute). Each univariate fitted model had the plot mean value of plant canopy height, temperature or NDVI as the dependent variable. For the explanatory variables, cultivar (i.e. genotype), replicate and block nested within replicate were fitted as random effects. The overall mean was the only fixed effect. Variance component estimates from each final model were used to estimate H^2 (Holland *et al.* 2003) as follows:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{r}} = \frac{\sigma_g^2}{\sigma_p^2}, \quad (3)$$

where σ_g^2 is the genotypic variance, σ_e^2 is the residual error variance and r is the number of replicates ($r=4$) in the field trial. The denominator of the equation is equal to the phenotypic variance, σ_p^2 . Standard errors of the heritability estimates were approximated using the delta method (Holland *et al.* 2003).

We assessed the validity of data obtained by the plant phenotyping system through comparisons with canopy temperature and NDVI data extracted from aerial images and plant height data collected by a calibrated bar-coded ruler from 900–1000 hours MST. We first calculated mean values for each plot ($n=200$) for these complementary data by averaging phenotypic measurements that had been screened for outliers with the statistical procedure described above. The degree of relationship (r^2) between phenotypes scored by two different methods was estimated using a simple linear model with PROC GLM in SAS ver. 9.3 (SAS Institute). This model included phenotype Method 1 as the dependent variable and phenotype Method 2 as the explanatory variable. Plant height data collected by the plant phenotyping system at 0700 hours MST on 21 July, and on 4 and 18 August at 0800 hours MST were used for the comparative analysis. Canopy temperature and NDVI

comparisons between aerial imagery (21 July at 1058 hours MST, 4 August at 1112 hours MST and 18 August at 1100 hours MST) and the plant phenotyping system (21 July at 1300 hours MST but with the canopy temperature data reprojected to 1058 hours MST, 4 August at 1100 MST hours and 18 August at 1100 hours MST) were also conducted. In addition, we compared the relationship between the canopy temperature data collected by the phenotyping system on 18 August at 1100 hours MST and the leaf RWC data collected from leaf discs sampled in the morning (0900–1030 hours MST) on 18 August.

Results

Field performance of the plant phenotyping system

The phenotyping system required an average of 0.32 h to traverse the complete set of plots (experimental plots, borders and alleys) that comprised the 0.27-ha field when measuring plant canopy height, reflectance and temperature, equivalent to an average rate of 0.84 ha h⁻¹ with an estimated field efficiency of ~77%. Thus, the system was in motion and continuously collecting phenotypic data for more than three-quarters of the time in the field. Slightly less than a quarter of the field time was attributed to the unproductive activities of turning and idle travel. The system recorded an average of 230.3 kB, 112.2 kB and 333.1 kB of canopy height, reflectance and temperature data, respectively, for a total of 675.6 kB. These values translate to 258 bytes of data per m², collected at a rate of 2.094 MB h⁻¹.

We found that the boom – the support structure for sensors – was very stable when the system was operating in the field. Plant growth regulators were not applied during the growing season in order to maximise phenotypic diversity, but even with the resulting tall and wide cotton plants, the system incurred minimal damage to plants within the trafficked rows due to its high clearance, as well as use of tyre shrouds and fenders. However, a notable limitation was that the ultrasonic proximity sensor only had a maximum clearance of 1.1 m (21 July) and 1.4 m (4 August and 18 August) when collecting data, which was not high enough to clear the tallest cotton plants completely. Notably, the maximum clearance for all sensors was not allowed to exceed 1.4 m because the field of view for the multispectral crop canopy and IRT sensors would have been negatively affected at higher elevations. Therefore, the inter-relationship of the optimal vertical height for multiple sensors needs to be further explored when simultaneously phenotyping several traits

on experimental plots that capture a wide range of height variation.

Phenotypic variability

We investigated the extent to which plant canopy height, NDVI, and temperature varied at different times on 3 days during the 2011 field season. On the first of the 3 days (21 July), moderate differences among cotton cultivars ($P \leq 0.0001$) were detected for canopy height and temperature (Table 2). Given that the water deficit treatment had been imposed on WL plots for less than 15 days, it is not surprising that treatment differences ($P > 0.05$) were not detected for height and NDVI – two traits that are associated with biomass. In contrast, water regime was significant for canopy temperature, a physiological trait that is expected to respond more rapidly to limiting soil moisture levels under high ambient temperatures. Time of day was also significant for canopy temperature, as the temperature increased by 29% and 37%, on average, from 0800 to 1300 hours MST for WW and WL plots, respectively (Fig. 2). On average, WL plots were ~5°C warmer than WW plots at 1300 hours MST under hot, sunny conditions, which is indicative of drought-stressed cotton plants having decreased leaf cooling capacity from a reduced transpiration rate (Carmo-Silva *et al.* 2012). Time of day was also significant for NDVI, with a 21% and 7% decrease in NDVI values from 0800 hours to 1300 hours MST for WL and WW plots, respectively. These results are consistent with higher leaf wilting for drought-stressed cotton plants from greater leaf dehydration as the day progresses (Carmo-Silva *et al.* 2012), thus altering leaf orientation, which resulted in higher soil exposure.

On the second and third phenotyping days (4 and 18 August), we identified larger differences among cotton cultivars ($P \leq 0.0001$) for canopy height, NDVI and temperature (Table 2). In addition, cultivars had a highly significant interaction ($P \leq 0.001$ and $P \leq 0.0001$) with the water regime for these three traits on both days. However, time of day and water regime were more highly significant ($P \leq 0.0001$) for these traits on 18 August, which can be explained by differences in the severity of the water deficit. On the second day (4 August), both WW and WL plots had received equal furrow irrigation amounts 6 days earlier in order to hydrate the field uniformly after problems with clogged drip irrigation tape and emitters were eliminated. As shown in Fig. 2, this less severe water deficit for WL plots resulted in weak water regime effects for canopy temperature

Table 2. Analysis of variance for canopy height, normalised difference vegetation index (NDVI) and temperature from 25 cotton cultivars at different times on 3 days in Maricopa, Arizona, in 2011

F-values and significance levels of fixed sources of variation are shown. *, significant at $P=0.05$; **, significant at $P=0.01$; ***, significant at $P=0.001$; ****, significant at $P=0.0001$; C, cultivar; TD, time of day; water regime, WR

Source	21 July			4 August			18 August		
	Height	NDVI	Temperature	Height	NDVI	Temperature	Height	NDVI	Temperature
C	34.48****	8.34	3.93****	85.28****	16.74****	7.96****	72.87****	36.54****	15.45****
TD	1.69	22.91***	224.45****	0.83	6.17**	52.39****	9.03***	15.91****	403.98****
WR	2.51	10.16	21.53***	54.97****	11.91**	2.56	84.62****	25.1****	67.19****
C × WR	1.81*	1.93	2.02**	3.56****	3.42****	2.56****	2.21***	2.61****	2.76****
TD × WR	0.55	3.79	0.1	0.18	0.56	0.27	0.65	5.68*	0.58
C × TD × WR	0.46	0.19	0.36	0.15	0.1	0.96	0.39	0.33	2****

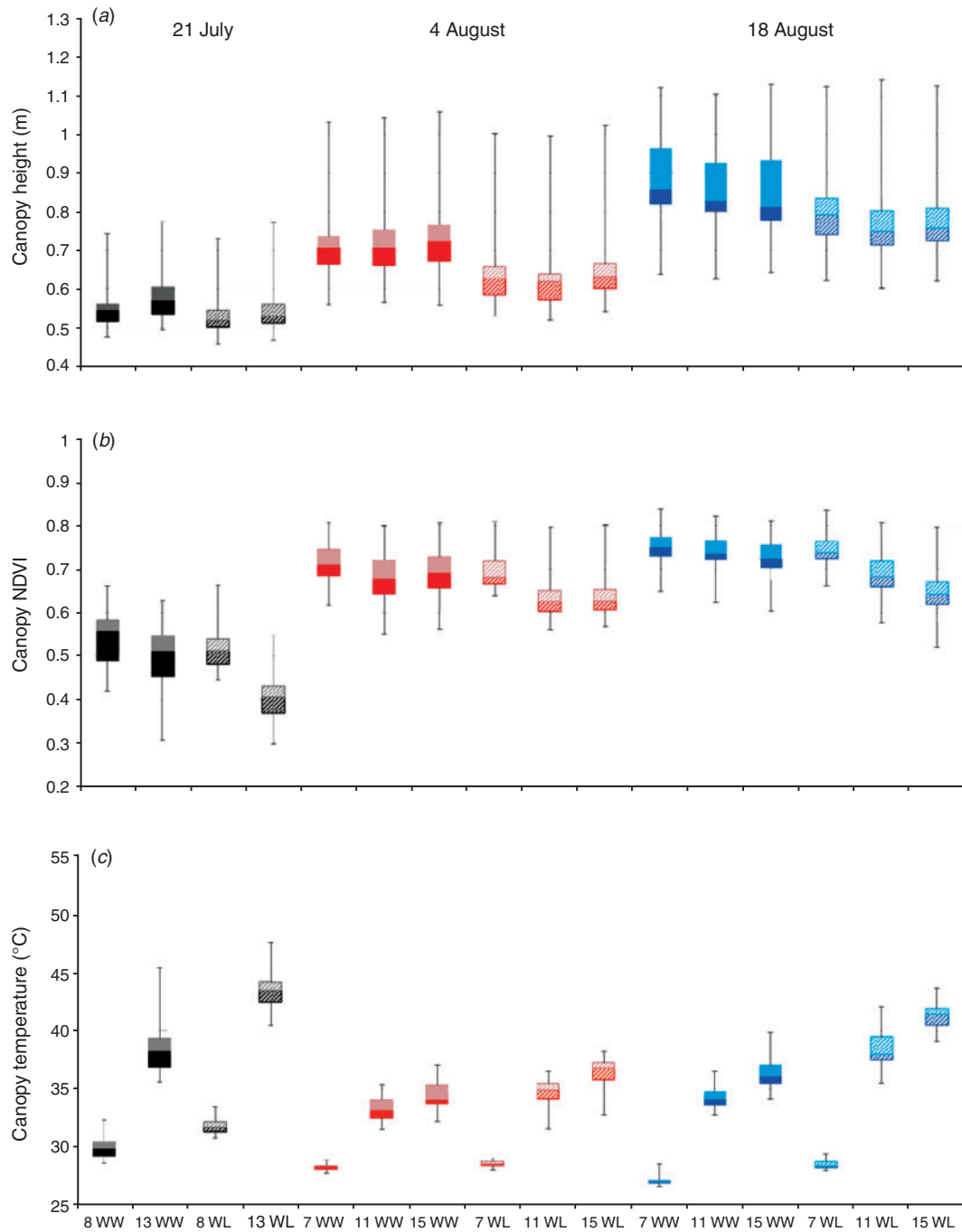


Fig. 2. Box-and-whisker plots of least-squares means for (a) canopy height, (b) normalised difference vegetation index (NDVI) and (c) temperature based on field evaluation of 25 Pima cotton cultivars at different time intervals (7, 0700–0800 hours; 8, 0800–0900 hours; 11, 1100–1200 hours; 13, 1300–1400 hours; 15, 1500–1600 hours Mountain Standard Time (MST)) for well watered (WW, solid boxes) and water-limited (WL, striped boxes) regimes on 21 July (shades of grey), 4 August (shades of red) and 18 August (shades of blue) in Maricopa, Arizona, in 2011. The bottoms and tops of the boxes are the first and third quartiles, respectively, whereas the colour separation line inside the box is the median. The ends of the whiskers represent the minimum and maximum of all the least-squares means.

($P > 0.05$) and NDVI ($P \leq 0.01$). The problem with the drip irrigation system had been resolved for 2 weeks when the cotton experimental plots were phenotyped on 18 August, thus allowing for a more controlled and severe water deficit in WL

plots. In agreement with our expectations, canopy temperature and NDVI values were more extreme for WL plots on 18 August relative to 4 August (Fig. 2). Unexpectedly, canopy height progressively decreased from 0700 hours to 1500 hours MST

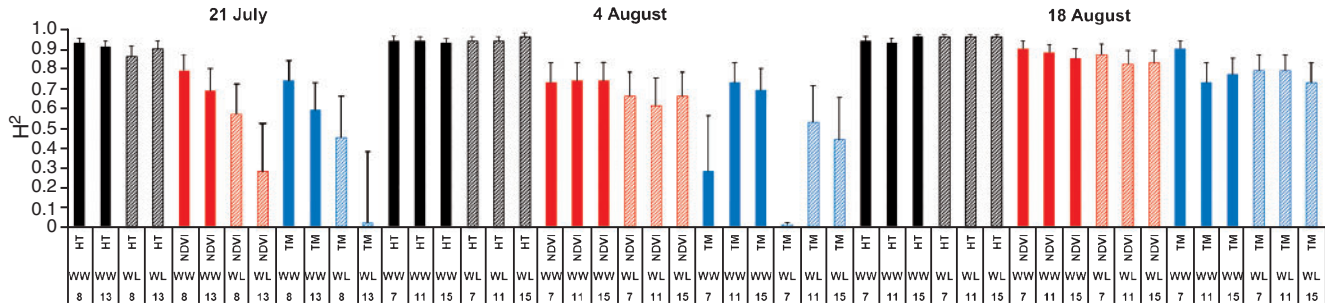


Fig. 3. Estimates of broad-sense heritability (H^2) on an entry-mean basis with their standard errors for canopy height (HT, black bars), normalised difference vegetation index (NDVI, red bars) and temperature (TM, blue bars) based on field evaluation of 25 Pima cotton cultivars at different time intervals (7, 0700–0800 hours; 8, 0800–0900 hours; 11, 1100–1200 hours; 13, 1300–1400 hours; 15, 1500–1600 hours Mountain Standard Time (MST)) for well watered (WW, solid bars) and water-limited (WL, striped bars) regimes on 21 July, 4 August and 18 August in Maricopa, Arizona, in 2011.

for all plots (time of day, $P < 0.0001$) on 18 August, which may have resulted from a high degree of leaf wilting on this particular day.

Broad-sense heritability (H^2) on an entry-mean basis corresponds to the level of phenotypic variability among cultivars that can be attributed to genetic effects (Holland *et al.* 2003), but in the context of our study of cultivars without estimated genetic relationships, H^2 will be conservatively interpreted as a measure of repeatability for a trait. The average broad-sense heritability on an entry-mean basis for canopy height was 0.94 for both WW and WL plots over the 3 days (Fig. 3). This value suggests that height was measured by our system with very high repeatability. On average, broad-sense heritability on an entry-mean basis for NDVI was 0.79 and 0.66 for WW and WL plots, respectively, for the 3 days. However, entry-mean heritabilities of 0.57 and 0.28 – the lowest for NDVI in this study – were estimated for WL plots at 0800 hours and 1300 hours MST on 21 July, respectively. Incomplete cotton canopy closure and vegetation gaps, in combination with intermittent clogging of the drip tape and emitters, were nongenetic sources of variation that probably contributed to the lower heritabilities for NDVI on this day. The wide range of entry-mean heritabilities for canopy temperature of WW and WL plots on 21 July ($H^2 = 0.02$ – 0.74) and 4 August ($H^2 = 0.01$ – 0.73) reflects the very high sensitivity of this trait to ineffective irrigation management practices. With a repaired drip irrigation system that delivered water more uniformly and consistently, however, estimates of broad-sense heritability on an entry-mean basis were larger for canopy temperature on 18 August, ranging from 0.73 to 0.90 and from 0.73 to 0.79 for WW and WL plots, respectively.

Evaluation of the plant phenotyping system

We evaluated the validity of the data obtained by the plant phenotyping system through comparisons with data collected by more expensive and laborious phenotyping methods. To validate plant canopy temperature and NDVI data collected by the phenotyping system, VNIR and TIR image data were collected using helicopter-mounted cameras. As shown in Table 3, helicopter-acquired canopy temperatures were closely related to temperature data from the nadir viewing IRTs, with discrepancies of less than 2°C . Considering the uncertainties in

Table 3. Leaf relative water content (RWC) of the Pima cotton cultivars Monseratt Sea Island (MS), Pima S-6 (S6), P62, 89590 and PSI425, under well watered (WW) and water-limited (WL) conditions

Two samples were taken from each of four replicated WW and WL plots on each of the indicated dates. Values are means \pm s.e. of the mean ($n = 8$)

Cultivar	Treatment	21 July RWC (%)	4 August RWC (%)	18 August RWC (%)
MS	WW	74.9 \pm 1.5	81.3 \pm 0.8	82.0 \pm 0.6
	WL	72.7 \pm 0.5	80.6 \pm 0.8	74.1 \pm 1.0
S6	WW	74.9 \pm 1.5	84.8 \pm 0.5	83.7 \pm 0.5
	WL	75.1 \pm 1.0	81.3 \pm 0.8	79.1 \pm 1.3
P62	WW	76.2 \pm 1.3	84.0 \pm 0.4	84.3 \pm 0.6
	WL	69.5 \pm 1.8	79.6 \pm 1.0	74.6 \pm 0.8
89590	WW	78.1 \pm 0.6	84.9 \pm 0.5	84.1 \pm 0.3
	WL	73.7 \pm 0.9	83.5 \pm 0.7	76.6 \pm 1.0
PSI425	WW	75.9 \pm 1.1	82.2 \pm 0.7	81.1 \pm 0.8
	WL	74.3 \pm 1.1	81.8 \pm 0.6	78.3 \pm 1.1

the FLIR camera calibration, atmospheric corrections and spatial resolution differences, we were still able to realise strong agreement ($r^2 = 0.75$ – 0.82) between the canopy temperature datasets. Incorporation of IRT data from the forward-pointing IRT would probably have improved the agreement by reducing effects from soil temperatures.

Canopy NDVI values from aerial images further confirmed the reasonableness of the ground-based observations of canopy NDVI, with a root mean square error of less than 0.10 for all 3 days. Agreement between NDVI data from the two sources was strongest for 4 and 18 August ($r^2 = 0.61$ – 0.62). These NDVI correlations indicate that reflectance data obtained by the phenotyping system were sensitive to plant canopy density changes in space and time. The exceptions to this assessment can be explained by instrumental differences and acquisition problems. NDVI values from the two systems resulted from different illumination and viewing angles, leading to significant discrepancies in apparent reflectances (Qi *et al.* 1997). The other correlation problem occurred only for 21 July, where the r^2 was 0.35. Considering the historical robustness of NDVI remote sensing, this relatively poor outcome was unusual but consistent with acquisition difficulties: images for this particular day had to be reconstructed from four separate

Table 4. Evaluation of canopy height, normalised difference vegetation index (NDVI) and temperature collected by the plant phenotyping system in 2011

Values shown are linear regression statistics relating data acquired from the plant phenotyping system to data collected manually (plant height) or via aircraft (NDVI and temperature). All r^2 values are significant at the $P=0.0001$ level. RMSE, root mean square error

Date	Height (m)		NDVI		Temperature (°C)	
	r^2	RMSE	r^2	RMSE	r^2	RMSE
21 July	0.76	0.04	0.35	0.09	0.82	1.92
4 August	0.79	0.06	0.62	0.05	–	–
18 August	0.78	0.07	0.60	0.04	0.75	1.58

aircraft passes due to problems with centring the helicopter directly over the field, which probably introduced radiometric errors from the compositing process.

We also assessed the relationship between phenotypic data collected by the plant phenotyping system and manual phenotyping approaches. Similar to canopy temperature comparisons, strong agreement ($r^2=0.76$ – 0.79) was detected between plant canopy height measured by the phenotyping system and that manually with a calibrated bar-coded ruler (Table 4). It is possible that these correlations would have been even stronger if two or more representative plants instead of one median plant were measured manually for each plot. In addition, it is possible that stronger agreements could have been obtained if all measured plants never exceeded the height of the sonar proximity sensors.

The RWC was determined in plants of five cotton cultivars (Table 3) as a measure of leaf water content. On 21 July, the RWC was below 70% for all cultivars, independent of the irrigation treatment, indicating that plants from both the WW and the WL plots were experiencing water deficit. The low RWC values were associated with the reduced availability of water caused by the clogged emitters in the drip irrigation lines. Regardless, plants from the WL plots, which had received the least amount of water, still had the lowest RWC values. On 4 August, the RWC was generally above 80%. On this date, all of the plants were well hydrated via application of flood irrigation 6 days before measurement. On 18 August, the RWC was higher in WW than in WL plants of each cultivar, demonstrating that previous issues with the drip irrigation system had been resolved and plants in the WL plots were experiencing higher water deficit compared with the WW plots. The strong inverse relationship ($r^2=0.74$) between the canopy temperature data collected by the plant phenotyping system and the RWC values was especially evident on 18 August (Fig. 4).

Discussion

The tractor-based phenotyping system proved capable of reliably acquiring and recording data for canopy temperature, height and reflectance on experimental plots of cotton plants throughout the growing season at much higher rates than can be achieved manually. Importantly, the phenotypic data collected were moderately to highly repeatable when management practices were optimal. In contrast to manual phenotyping protocols that

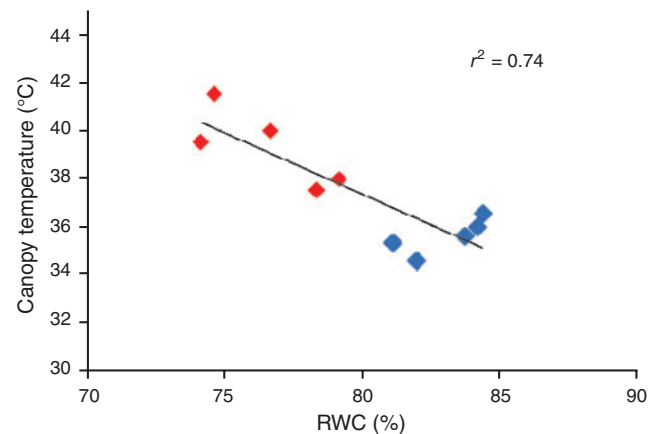


Fig. 4. The relationship between canopy temperature and leaf relative water content (RWC) for the Pima cotton cultivars Monseratt Sea Island, Pima S-6, P62, 89590 and PSI425, under well watered (WW) and water-limited (WL) conditions on 18 July 2011. Canopy temperature data were collected by the plant phenotyping system (i.e. tractor-mounted infrared radiometer) at 1100 hours Mountain Standard Time (MST) and the RWC values were based on samples collected from 0900 hours to 1030 hours MST. The WW and WL plots are demarcated in blue and red font colours, respectively.

require visual selection of a few representative plants, the suite of georeferenced sensors permitted unbiased collection of phenotypic data from many plants within a plot. As a powerful complement, our development of a custom open-source software package to post-process the collected phenotypic data efficiently helped to alleviate another significant bottleneck in high-throughput phenotyping. The strong agreement of these phenotypic data with the values obtained from aerial images and manual measurements further supports the conclusion that the plant phenotyping system is well suited for field-based high-throughput phenotyping of cotton.

Even though the phenotyping system was effective for collecting data from plants in the field, there are opportunities to improve the system further. As an example, the data acquisition rate of the platform could be enhanced through several complementary approaches. The number of rows monitored could be increased by extending the booms laterally and adding additional sets of sensors. The downsides to this change, however, are increased platform complexity and a greater likelihood of inconsistencies among sensor calibrations, as well as data logger connection problems. In addition, if plots are arranged in longer runs, the number of turnarounds will be substantially reduced, saving time and increasing field efficiency.

The spatial resolution of the sensor data was limited by the speed of operation of the platform through the field and by the sampling frequency of the data collection system. Data loggers like the ones used in this platform were designed to perform at very high sampling frequencies (100 Hz) but the response time of individual sensors was relatively slow. IRT and sonar transducers did not update output signals faster than 2 Hz. Moreover, current GPS receivers output serial data at a maximum frequency of 10 Hz, but logging and parsing data from more than one GPS string slowed down the data acquisition process significantly. Unifying the data collection into a single logging system would

Table 5. Comparison of four options for field-based, high-throughput phenotyping

Evaluations are subjective assessments that draw from multiple information sources. VL, very low; L, low; Mo, moderate; Me, medium; H, high; VH, very high

Attribute or issue	Hand-held	Cart	Tractor	Unmanned aerial system
Vehicle cost	L	L	H	H
Cost of operation exclusive of operator	Mo	Mo	Mo	Mo
Operator training requirements	L	L	Mo	H
Ease of manoeuvring the vehicle among field plots	H	Mo	L	H
Ease of controlling the height of sensors	Mo	H	H	Mo
Ease of simultaneously deploying multiple sensors	L	H	H	L
Ease of simultaneously acquiring data over different rows or plots	L	Mo	H	H
Ease of providing power for active sensors	L	Mo	H	L
Ease of stop-and-go operation	H	H	Mo	Mo
Maximum clearance over crop	H	Mo	Mo	VH
Potential for the vehicle to interfere with reflectance or thermometric readings	L	L	Mo	VL
Ease of acquiring data when soil is wet	Me	L	L	H
Ease of continuously acquiring data under high temperatures	L	L	H	Me
Ease of acquiring data under high winds	Me	Me	H	L
Risk of causing soil compaction	L	L	Mo	None
Risk of damaging plants in a closed canopy	L	Mo	H	VL
Ease of adjustment for different row spacing	H	Me	L	H
Ease of transportation to different locations	H	Me	Me	H

simplify deployment, increase reliability and help standardise data collection protocols. Ideally, individual instruments should output signals that have meaningful values at frequencies up to 20 Hz and they should require power in a standard format (e.g. 12 V direct current).

Future improvements in electronics and signal processing will produce sensor technology and data acquisition systems permitting higher throughput. As an example, a platform traveling at 1 m s^{-1} and logging sensor and position data at 20 Hz would return data points every 5 cm. This will be important to the application of field-level high-throughput phenotyping because increasing the data acquisition rate may allow for better discrimination of within-plot variation. As pointed out by Scotford and Miller (2004), the coefficient of variation for height in a different crop, wheat (*Triticum aestivum* L.), is related to fine-scale variations in tiller number. Our observations suggest that within-plot variation in cotton canopy temperature is partially related to variation in interplant spacing, and statistical methods are needed to remove this nongenetic source of variation in the postprocessing of canopy temperature data.

We recognise, too, that tractor-based systems have limitations. Table 5 presents a comparison with alternatives, including hand-held sensors, manual powered carts, tractors and unmanned aerial systems. The most serious issues are inherent to entry of any wheeled vehicle into research field plots. One option to overcome such a problem is to restrict vehicles to all-weather traffic lanes along the edge of plots and use a lateral boom to reach over plots. This strategy would be especially effective if multiple types of sensors were combined in a single light-weight instrument head. On the plus side, tractor-based systems provide great flexibility in terms of mounting and powering devices. They are built on well understood technologies that are familiar to most research farms and they are deployable for a wide range of field situations.

As acquisition of raw data becomes less of a constraint, more attention will be need to be paid to other aspects of field-based high-throughput phenotyping. One issue relates to temporal scales. Most proximal sensing appears to have emphasised time series over the crop growing season. Greater insights into dynamic physiological responses to water deficits or acute heat stress may require multiple measurements over a diurnal cycle, including night-time observations, to track potential recovery. A second issue relates to analytical approaches that maximise the information content obtained from multiple sensors. Scotford and Miller (2004) showed that combining canopy reflectance and height data allowed for a more accurate estimation of LAI in wheat. White *et al.* (2012) proposed using inverse modelling to estimate more biologically fundamental traits such as stem elongation rates or stomatal responses to humidity. Other options include analysing the time series as repeated measures, functional growth curves analysis (Hunt 1979) and using multivariate statistical methods.

Summary and Perspectives

We showed that the plant phenotyping system can be used to phenotype plants within specific time intervals needed to capture the rapid response of plants to environmental stress with high repeatability under optimal management conditions. This ability provides a powerful new tool for plant breeding programmes that seek to develop high-yielding cultivars with enhanced adaptation to stress-prone environments. In cotton and many other crop species, phenotypic traits are regularly measured at a fixed time point that is typically at or near plant maturity. With such an approach, it is only possible to estimate the accumulated effects of QTL from the beginning of plant development to the fixed time of observation (Wu *et al.* 1999). Our plant phenotyping system provides an opportunity to measure traits repeatedly throughout plant development, facilitating the identification of

QTL that could be differentially expressed throughout the developmental program of a trait. The next phase for our field-based phenotyping systems will be the integration of high-resolution imaging and spectral technologies to measure an even wider diversity of traits across larger plant populations.

Acknowledgements

This research was supported by Cotton Inc. (PA-S and MAG) and United States Department of Agriculture – Agricultural Research Service (USDA-ARS) (MAG, KRT, ANF, MES and JWW). Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA. The USDA is an equal opportunity provider and employer. We especially thank Doug Hunsaker for consultation on irrigation scheduling, as well as Kristen Cox, Bill Luckett, Joel Gilley and Bob Strand for providing excellent technical expertise. PA-S and MAG contributed equally to this work.

References

- Allen RG, Pereira LS, Raes D, Smith M (1998) 'Crop evapotranspiration – guide-lines for computing crop water requirements. FAO irrigation and drainage paper 56.' (Food and Agriculture Organization of the United Nations: Rome)
- Box GEP, Cox DR (1964) An analysis of transformations. *Journal of the Royal Statistical Society. Series B. Methodological* **26**, 211–252.
- Brown PW (1989) 'Accessing the Arizona Meteorological Network (AZMET) by vomputer. Extension report no. 8733.' (University of Arizona: Tucson)
- Carmo-Silva AE, Gore MA, Andrade-Sanchez P, French AN, Hunsaker DJ, Salvucci ME (2012) Decreased CO₂ availability and inactivation of Rubisco limit photosynthesis in cotton plants under heat and drought stress in the field. *Environmental and Experimental Botany* **83**, 1–11. doi:10.1016/j.envexpbot.2012.04.001
- Comar A, Burger P, de Solan B, Baret F, Daumard F, Hanocq JF (2012) A semi-automatic system for high throughput phenotyping wheat cultivars in-field conditions: description and first results. *Functional Plant Biology* **39**, 914–924. doi:10.1071/FP12065
- Furbank RT, Tester M (2011) Phenomics – technologies to relieve the phenotyping bottleneck. *Trends in Plant Science* **16**, 635–644. doi:10.1016/j.tplants.2011.09.005
- Fussell J, Rundquist D, Harrington JA (1986) On defining remote sensing. *Photogrammetric Engineering and Remote Sensing* **52**, 1507–1511.
- Göttsche F-M, Olesen FS (2001) Modelling of diurnal cycles of brightness temperature extracted from METEOSAT data. *Remote Sensing of Environment* **76**, 337–348. doi:10.1016/S0034-4257(00)00214-5
- Holland JB, Nyquist WE, Cervantes-Martínez CT (2003) Estimating and interpreting heritability for plant breeding: an update. *Plant Breeding Reviews* **22**, 9–112.
- Houle D, Govindaraju DR, Omholt S (2010) Phenomics: the next challenge. *Nature Reviews. Genetics* **11**, 855–866. doi:10.1038/nrg2897
- Hunt R (1979) Plant growth analysis: the rationale behind the use of the fitted mathematical function. *Annals of Botany* **43**, 245–249.
- Kimes D (1981) Remote sensing of temperature profiles in vegetation canopies using multiple view angles and inversion techniques. *IEEE Transactions on Geoscience and Remote Sensing* **GE-19**, 85–90. doi:10.1109/TGRS.1981.350357
- Kutner MH, Nachtsheim CJ, Neter J, Li W (2004) 'Applied linear statistical models.' 4th edn. (McGraw-Hill: Boston)
- Lan Y, Zhang H, Lacey R, Hoffman W, Wu W (2009) Development of an integration sensor and instrumentation system for measuring crop conditions. *Agricultural Engineering International: CIGR Journal* **11**, 1–16.
- McCarthy C, Hancock N, Raine S (2010) Apparatus and infield evaluations of a prototype machine vision system for cotton plant internode length measurement. *Journal of Cotton Science* **14**, 221–232.
- Milton EJ (1987) Principles of field spectroscopy. *International Journal of Remote Sensing* **8**, 1807–1827. doi:10.1080/01431168708954818
- Montes JM, Melchinger AE, Reif JC (2007) Novel throughput phenotyping platforms in plant genetic studies. *Trends in Plant Science* **12**, 433–436. doi:10.1016/j.tplants.2007.08.006
- Piepho H-P, Möhring J (2007) Computing heritability and selection response from unbalanced plant breeding trials. *Genetics* **177**, 1881–1888. doi:10.1534/genetics.107.074229
- Qi J, Pinter P, Clarke TR, Kimball BA, Moran MS (1997) Diagnostic assessments of plant condition using multiangular remote sensing measurements and BRDF models. *Journal of Remote Sensing* **1**, 25–29.
- Ruixiu S, Wilkerson JB, Wilhelm LR, Tompkins FD (1989) A microcomputer-based morphometer for bush-type plants. *Computers and Electronics in Agriculture* **4**, 43–58. doi:10.1016/0168-1699(89)90013-6
- Rundquist D, Perk R, Leavitt B, Keydan G, Gitelson A (2004) Collecting spectral data over cropland vegetation using machine-positioning versus hand-positioning of the sensor. *Computers and Electronics in Agriculture* **43**, 173–178. doi:10.1016/j.compag.2003.11.002
- Scottford IM, Miller PCH (2004) Estimating tiller density and leaf area index of winter wheat using spectral reflectance and ultrasonic sensing techniques. *Biosystems Engineering* **89**, 395–408. doi:10.1016/j.biosystemseng.2004.08.019
- White JW, Conley MM (2013) A flexible, low-cost cart for proximal sensing. *Crop Science* **53**, 1646–1649. doi:10.2135/cropsci2013.01.0054
- White JW, Andrade-Sanchez P, Gore MA, Bronson KF, Coffelt TA, Conley MM, Feldmann KA, French AN, Heun JT, Hunsaker DJ, Jenks MA, Kimball BA, Roth RL, Strand RJ, Thorp KR, Wall GW, Wang G (2012) Field-based phenomics for plant genetics research. *Field Crops Research* **133**, 101–112. doi:10.1016/j.fcr.2012.04.003
- Wu WR, Li WM, Tang DZ, Lu HR, Worland AJ (1999) Time-related mapping of quantitative trait loci underlying tiller number in rice. *Genetics* **151**, 297–303.