Utilization of a high-throughput shoot imaging system to examine the dynamic phenotypic responses of a C4 cereal crop plant to nitrogen and water deficiency over time E.H. Neilson, A.E. Edwards, C.K. Blomstedt, B. Berger, B. Lindberg Møller, R.M. Gleadow

## **SUPPLEMENTARY INFORMATION**

## **SUPPLEMENTARY FIGURES**



Adapted from Archontoulis and Miguez (2013)

Supplementary Figure 1 Schematic overview for model selection and analysis (Abbreviations: 3PL, three parameter logistic; 4PL, four parameter logistic; AIC, Akaike information criterion). The models chosen for analysis are underlined.

Archontoulis SV, Miguez FE. 2013. Nonlinear Regression Models and Applications in Agricultural Research. Agronomy Journal 105, 1-13. Reprinted by Permission, ASA, CSSA, SSSA.



**Supplementary Figure 2** Relationship between shoot biomass and projected leaf area calculated from image analysis for Sb and HyA plants subject to high and low water conditions. Projected leaf area provides a strong indication of plant growth with a significant, positive correlation observed (P < 0.0001;  $R^2$  = 0.97;  $y$  = 78.0x - 2.71).



**Supplementary Figure 3** Assessment of Relative Growth Rate (RGR) derived from the power law and 3PL models selected for the nutrient (A-C) and water limiting experiments (D-F), respectively. RGR(imaged) was derived from the models using continuous data, compared to RGR(biomass) calculated from initial biomass of a sample of plants and final biomass of replicate plants (A and D); RGR(area) was calculated using the final leaf area of individual plants and the pooled mean initial leaf area (B and E); and RGR(individual) was where both the initial and final leaf area of individual plants were used to calculate the growth. A strong correlation was detected between RGR(biomass) and RGR(imaged) when plants grown under different nitrogen levels (A, B;  $R^2$  = 0.93 and 0.91, respectively) but not when plants were grown under water-limited conditions (D, E;  $R^2$  = 0.83 and 0.88, respectively). A very high correlation was detected in both nutrient and water limiting experiments when the initial leaf area for each individual plant was used to calculate RGR (C, D;  $R^2$ =0.93 and 94, respectively).



**Supplementary Figure 4.** Phenotypic response in a representative *Sorghum bicolor* (Sb) plant subjected to limited watering over four weeks as observed in images taken from above (A) and from the side (B). Leaf area (grey symbols, C) initially increased but then decreased as a result of leaf curling and senescence (white symbols).



**Supplementary Figure 5:** Effect of seedling size (small, intermediate and large) on HyA (A-C) and Sb (D-F) growth under high and low watering regimes. Calculated projected leaf area (A, D), absolute growth rate (B, E) and relative growth rate (C, F) for the asymptotic 'waterlimited' dataset fitted with the three parameter logistic model.



**Supplementary Figure 6:** Differences in NIR reflectance (A), moisture content (B) and LMA (leaf mass per area) (C) in Sb and HyA plants subjected to well watered and water limiting conditions. The significant differences in NIR (A) could not be accounted for by moisture content alone, as demonstrated by the similarity in moisture content observed between the two sorghum lines under low water conditions (B). However, a significant difference was observed, in LMA, which is used as a proxy for leaf thickness (C).

## **SUPPEMENTARY TABLES**

## **Table S1**

Summary of the two experiments including cultivars and the dates of sowing and imaging. The first experiment was a nitrogen trial using two different forage sorghum hybrids (HyA and Sweet Jumbo -LPA). The second experiment was a water supply trial using one forage sorghum hybrid (HyA) and a variety of grain sorghum (*Sorghum bicolor*; Sb).



# **Table S2**

Summary of plant traits derived using LemnaGrid software using phenotypic imaging recorded in The Plant Accelerator®.



# **Table S3 (A)**

Growth curve parameters for the modelled projected shoot area vs days after sowing for the nitrogen experiment with the calculated  $R^2$  values (See equation 2 and equation 3 in the Materials and Methods of the main text). Two varieties of forage sorghum (HyA and HyB) were grown at three different levels of nitrogen for 4 weeks and imaged daily.



# **Table S3 (B)**

Growth curve parameters for the modelled projected shoot area vs days after sowing for the watering experiment with the calculated  $R^2$  values (See equation 2 and equation 3 in the Materials and Methods of the main text).). Two varieties of sorghum, a hybrid (HyA) and a variety of *Sorghum bicolor* (Sb) were grown for 4 weeks and either well-watered (High W) or watered to 25% of field capacity (Low W).



## Table S3 (C)

Equations for eight models used to assess the growth of sorghum in two sets of data obtained from The Plant Accelerator ®. Growth models can be classified under two broad categories, those that assume no asymptotic final size exists (e.g. exponential growth) and those that do assume one exists (eg logistic curve). Non-Asymptotic growth models make the assumption that growth will continue indefinitely. Although this assumption may be unrealistic for plants through the entire lifetime of a plant, they may be appropriate for modelling the growth of young seedlings. The ability of each model to describe the data was assessed using convergence and correlative analysis (see Table S3A and S3B).

#### **Exponential**

The exponential model implies that the rate of change of a measured quantity is proportional to the measured quantity.

$$
M(t) = M_0 e^{rt}
$$
  
\n
$$
AGR = \frac{dM}{dt} = M_0 r e^{rt}
$$
  
\n
$$
RGR = \frac{1}{M} \frac{dM}{dt} = r
$$
  
\nEquation A.2  
\nEquation A.3

Where M is the measured value t is the time after sowing  $M_0$  is the initial measured value at time  $t_0$ r is a parameter controlling the relative growth rate.

#### **Power Law**

The power law model is a non-asymptotic model that allows both AGR and RGR to vary throughout time.  $0 > \beta > 1$  implies a progressive decrease in Relative growth rate. The linear model is obtained when  $\beta$ =0 and the model converges to the exponential model as  $\beta$  approaches 1.

$$
M(t) = (M_0^{1-\beta} + rt(1-\beta))^{\frac{1}{1-\beta}}
$$
 Equation B.1  
\n
$$
AGR = \frac{dM}{dt} = r (M_0^{1-\beta} + rt(1-\beta))^{\frac{\beta}{1-\beta}}
$$
 Equation B.2  
\n
$$
RGR = \frac{1}{M} \frac{dM}{dt} = \frac{r}{(M_0^{1-\beta} + rt(1-\beta))}
$$
 Equation B.3

Where

M is the measured value t is the time after sowing  $M_0$  is the initial measured value at time  $t_0$ r is a parameter controlling the growth rate.  $\beta$  controls the progressive changes to RGR.

#### 3-paramater logistic

The 3-parameter logistic model is the most commonly used asymptotic model used to model growth. The lower asymptote of the 3 parameter version of the logistic curve is fixed at 0. Although this may not be strictly realistic as there is an inherit seed weight associated with a plant, this is generally small enough to be ignored for the purpose of curve fitting. In addition the inflection point of the curve (point at which the AGR is maximised) falls at half the expected upper asymptotic value. Other versions of the logistic curve can relax one or both of these constraints. The 3-parameter logistic function is implemented in R with the SSlogistic function in the nlme package.

$$
M(t) = \frac{M_{max}}{1 + e^{\left(\frac{t_{mid} - t}{k}\right)}}
$$
 Equation C.1  

$$
AGR = \frac{dM}{dt} = \frac{M_{max}e^{\frac{t_{mid} + t}{k}}}{k\left(e^{\frac{t_{mid}}{k} + e^{\frac{t}{k}}\right)^{2}}}
$$
 Equation C.2  

$$
RGR = \frac{1}{M}\frac{dM}{dt} = \frac{e^{\frac{t_{mid} + t}{k}\left(1 + e^{\frac{t_{mid} - t}{k}}\right)}}{k\left(e^{\frac{t_{mid}}{k} + e^{\frac{t}{k}}\right)^{2}}}
$$
 Equation C.3

Where M is the measured value t is the time after sowing  $M_{\text{max}}$  is the upper asymptote t<sub>mid</sub> represents the inflection point of the model k is a parameter controlling the maximal growth rate.

## **Gompertz**

The gompertz model allows RGR to fall exponentially over time. This differs from the logistic model in that the time of maximal growth appears at approximately 37% of the final expected biomass. Accordingly, the end result of the upper asymptote is approached more slowly than the lower asymptote. Like the 3 parameter logistic model, the gompertz model assumes an initial measured value of 0. The gompertz model is implemented in R with the SSgompertz function in the nlme package.

$$
M(t) = Asym \, e^{-kr^{t}}
$$
  
\n
$$
AGR = \frac{dM}{dt} = k \, Asym \, r^{t} \left( -e^{-kr^{t}} \right) \log_{e} r
$$
  
\n
$$
RGR = \frac{1}{M} \frac{dM}{dt} = -kr^{t} \log_{e} r
$$
  
\nEquation D.2  
\nEquation D.3  
\nWhere

M is the measured value t is the time after sowing Asym is the upper asymptote k displaces the graph along the y axis r controls the maximal growth rate

### **Beta Function**

The beta function is a 3 parameter sigmoidal function that allows flexibility in the time of maximal AGR. It does this using only 3 parameters allowing for easier convergence. However the model is not continuous (Yin et al, 2003).

$$
M(t) = \begin{cases} 0, & t < 0\\ M_{max} \left( 1 + \frac{t_{end} - t}{t_{end} - t_{mid}} \right) \left( \frac{t}{t_{end}} \right)^{\frac{t_{end}}{(t_{end} - t_{mid}}}, \quad 0 \le t \le t_{end}
$$
 Equation E.1  
 
$$
M_{max}, \quad t > t_{end}
$$

$$
AGR = \frac{dM}{dt} = \begin{cases} 0, & t < 0\\ c_m \left(\frac{t_{end} - t}{t_{end} - t_{mid}} \right) \left(\frac{t}{t_{mid}}\right)^{\frac{t_{mid}}{t_{end} - t_{mid}}, & 0 \le t \le t_{end} \quad \text{Equation}
$$

Where

$$
c_m = M_{max} \left( \frac{2t_{end} - t_{mid}}{t_{end}(t_{end} - t_{mid})} \right) \left( \frac{t_{mid}}{t_{end}} \right)^{t_{end} - t_{mid}}
$$
  
\n
$$
RGR = \frac{1}{M} \frac{dM}{dt} = \frac{(2t_{end} - t_{mid})(t_{end} - t_{mid}) (t_{end} - t_{
$$

Where M is the measured value t is the time after sowing  $M_{\text{max}}$  is the maximal measured value  $t_{end}$  is the time that  $M_{max}$  is achieved t<sub>mid</sub> represents the inflection point of the model

### 4-paramater Logistic

The 4-parameter version of the logistic function relaxes the constraint on the lower asymptote, allowing non-zero values. The 3-parameter logistic model is obtained when M<sub>min</sub> equals 0. The four parameter logistic function is implemented in R with the SSfpl function in the nlme package.

$$
M(t) = \frac{M_{min}(M_{max} - M_{min})}{1 + e^{\frac{\left(\frac{t_{mid} - t}{k}\right)}}}
$$
Equation  

$$
AGR = \frac{dM}{dt} = \frac{M_{min}(M_{max} - M_{min})e^{\frac{t_{mid} - t}{k}}}{k\left(e^{\frac{t_{mid} - t}{k}} + 1\right)^{2}}
$$
Equation  

$$
RGR = \frac{1}{M}\frac{dM}{dt} = \frac{1}{k\left(e^{\frac{t - t_{mid}}{k}} + 1\right)}
$$
Equation

Where M is the measured value t is the time after sowing  $M_{min}$  is the lower asymptote  $M_{\text{max}}$  is the upper asymptote t<sub>mid</sub> represents the inflection point of the model  $n F.1$ 

 $F.2$ 

 $n F.2$ 

 $nF.3$ 

k is a parameter controlling the maximal growth rate.

*Weibull*

$$
M(t) = \frac{M_{max} - drop}{e^{e^{r}t^{p}}}
$$
  
\n
$$
AGR = \frac{dM}{dt} = p(drop - M)t^{p-1}e^{r-e^{r}t^{p}}
$$
  
\n
$$
RGR = \frac{1}{M}\frac{dM}{dt} = p(-e^{r})t^{p-1}
$$
  
\nEquation G.3  
\nEquation G.3

Where M is the measured value t is the time after sowing  $M<sub>max</sub>$  is the upper asymptote drop represents the change from  $M_{\text{max}}$  to the y intercept r controls the maximal growth rate p is a numerical constant representing the power to which x is raised

#### *Richards*

The Richards curve is a 4 parameter version of the logistic function that relaxes the constraint on the location of maximal growth. The 3-parameter logistic model is obtained when p equals 0. The Richards function is implemented in R with the *SSRichards* function in the NRAIA package.

$$
M(t) = \frac{M_{max}}{\left(1 + e^{\frac{t_{mid} - t}{k}}\right)^{e^{-p}}}
$$

$$
AGR = M_{max}e^{-\frac{kp - t_{mid} - t}{k}}\left(e^{\frac{t_{mid} - t}{k}} + 1\right)^{-e^{-p} - 1}
$$

Equation H.1

Equation H.2

M is the measured value t is the time after sowing  $M_{min}$  is the lower asymptote  $M<sub>max</sub>$  is the upper asymptote  $t_{mid}$  represents the inflection point of the model k is a parameter controlling the maximal growth rate. p controls the location of maximal growth.

## **Table S4**

R scripts used to calculate the fitted growth models, absolute growth rates (AGR) and relative growth rates (RGR) along with the 95% confidence intervals for sorghum varieties used in the nitrogen and water-limiting experiments. Models are adapted from (Paine *et al.*, 2012). Text beginning with '#' contains comments explaining the function of the code following the statement, how to use the function and what it produces. The main driver of the script (output.growth.model) also shows a brief "how to use" with example data that is included in R.

```
# Summarizer function which takes the data set and an alpha value and 
returns
# the high and low value for each data point within the confidence interval
specified
# in alpha (Adapted from Paine et al 2012)
#
# Example usage
# dataMatrix <- matrix(c(0,1,2,3,4,5,6,7,8,9,10))
# summarizer(list(growth=dataMatrix), 0.05)
#
# Expected Result:
# growth.lo growth.hi
\# 0.5 9.5
summarizer <- function(dat, alpha){
   n <- length(dat)
   quantiles <- c(alpha/2, 1-(alpha/2))
   CIs <- data.frame(matrix(NA, ncol(dat[[1]]), n*2))
   names(CIs) <- paste(rep(names(dat), each = 2), c("lo", "hi"), sep = ".")
  for(i in 1:n){
   CIs[,(2*i-1):(2*i) <- t(apply(dat[[i]], 2, quantile, quantiles, na.rm
= T)
  }
   return(CIs)
}
```
et al 2012) # This cfunction can be used to produce values for the model provided, its first # derivitive (Absolute growth rate) and the first derivitive devided by the model # (Relative growth rate) and optionally the confidence intervals around the model.  $#$ # The function takes any nls or gnls object (from the 'nlme' R package). The function # also requires the underlying model function to generate the first derivitive of the # model. The function will also optionally calculate the confidence intervals for the # growth model, the AGR and the RGR with the provided alpha value. (Adapted from Paine et al 2012) # # Parameters:

# General function to generate output data for models (Adapted from Paine

# fit: a nls or gnls fitted model (from the nlme package)

# times: a sequence of time points to calculate

# formula: an expression containing the formula for the model used to fit the data

# indVar: string representation of the indepedant variable used in the formula provided (defaults to 'x')

# CI: Boolean value determining whether or not to calculate the confidence intervals (defaults to False)

# alpha: the alpha cut-off for confidence intervals (defaults to 0.05)

#

# Output:

# params: A list of the calculated parameters for the fitted model

# stats: A list containing the R^2, AIC and RMSE values

# data: The fitted and rediduals calculated from the model

# rates: The M (model value), AGR, RGR and optionally the high and low confidence intervals

# for the M, AGR and RGR for each of the time points provided

S16

```
# summary: Summary statistics including:
# AvAGR: Average AGR
# AvRGR: Average RGR
# PeekTime: The time of maximal absolute growth rate
# PeekAGR: The maximal absolute growth rate
# PeekTime.hi: The high confidence interval for the time 
of maximal absolute growth rate
# PeekTime.lo: The low confidence interval for the time of 
maximal absolute growth rate
# PeekAGR.hi: The high confidence interval for maximal 
absolute growth rate
# PeekAGR.lo: The low confidence interval for maximal 
absolute growth rate
#
# Example usage (fitting the R default ChickWeight dataset to the 3 
parameter logistic model)
# library(nlme)
# logisticFit <- gnls(weight ~ SSlogis(Time, Asym, xmid, scal), 
ChickWeight, weights=varPower())
# logisticFit <- gnls(weight ~ SSlogis(Time, Asym, xmid, scal), 
ChickWeight, weights=varPower())
# formula <- expression(Asym/(1+exp((xmid-x)/scal)))
# times \langle -\ \text{seq}(1, 50, b) = 0.1 \rangle# model <- output.growth.model(logisticFit, times, formula, 
indVar='x', CI=T, alpha=0.05)
#
library(mvtnorm)
output.growth.model <- function(fit, times, formula, indVar='x', CI=F, 
alpha=0.05){
  param <- coef(fit)
  #
  if(inherits(fit, "gnls")){
    fitted <-fit$fitted
    resid <- fit$resid
  }else{ #Assume nls
```

```
 fitted <- fit$m$fitted()
   resid <- fit$m$resid()
 }
data \leftarrow data .frame (fitted = fitted, resid = resid) mss <- sum((fitted - mean(fitted))^2)
 rss <- sum(resid^2)
R2 \leq -mss/(mss + rss) rmse <- sqrt(rss)
AIC \leftarrow AIC(fit)
 assign(toString(indVar), times)
 AGRFormula <- D(formula,toString(indVar))
 rates = data.frame(
   times = times,
   M = eval(formula, as.list(param)),
   AGR = eval(AGRFormula, as.list(param))
\overline{\phantom{0}}rates$RGR = rates$AGR/rates$M
 peekAGRm <- max(rates$AGR)
 peekAGRt <- rates$times[which(rates$AGR==max(rates$AGR))]
 avAGR <- mean(rates$AGR)
 avRGR <- mean(rates$RGR)
if(CI == T) if(inherits(fit, "gnls")){
     cov <- fit$varBeta
   }else{
     cov <- summary(fit)$cov
   }
   .x <- data.frame(rmvnorm(n=1000, mean=param, sigma=cov))
  M \leq -AGR \leq -RGR \leq -matrix(NA, ncol = length(time), nrow = nrow(.x)) peakAGRt <- c()
```

```
 peakAGRm <- c()
    for(i in 1:nrow(.x))\{ M[i,] <- eval(formula, as.list(.x[i,]))
       AGR[i, ] <- eval(AGRFormula, as.list(.x[i,]))
      RGR[i, ] \leftarrow AGR[i, ]/M[i, ] peakAGRm[i] <- max(AGR[i, ])
       peakAGRt[i] <- rates$times[which(AGR[i, ]==max(AGR[i, ]))]
     }
    CIs \leftarrow \text{summarizer}(\text{list}(M = M, \text{ AGR} = \text{AGR}, \text{RGR} = \text{RGR}), \text{alpha}) rates = cbind(rates, CIs)
    out \langle- list(params = param, stats = c(R2 = R2, AIC = AIC, RMSE = rmse),
                  data = data, rates = rates, summary=c(AvAGR=avAGR, AvRGR=avRGR,
                              PeekTime=peekAGRt, PeekAGR=peekAGRm,
                              PeekTime.hi = quantile(peakAGRt, 1-alpha/2, na.rm 
= T)[1],
                              PeekTime.lo = quantile(peakAGRt, alpha/2, na.rm = 
T)[1],
                              PeekAGR.hi = quantile(peakAGRm, 1-alpha/2, na.rm 
= T)[1], PeekAGR.lo = quantile(peakAGRm, alpha/2, na.rm = 
T)[1]
                   ))
```
}else{

```
out \langle- list(params = param, stats = c(R2 = R2, AIC = AIC, RMSE = rmse),
            data = data, rates = rates, summary=c(AvAGR=avAGR, AvRGR=avRGR,
                        PeekTime=peekAGRt, PeekAGR=peekAGRm))
```
}

}

Table S5: Chlorophyll (Chl), HCN (dhurrin), nitrate (NO<sub>3</sub>) concentration, and total nitrogen allocation in the leaf tissue of two hybrid sorghum varieties grown under different nitrogen conditions. Mean ± standard error and significance of a two-way ANOVA are presented for each variety (V) and treatment (T).



\*P< 0.05; \*\*P< 0.01; \*\*\*P< 0.001; ns, not significant

Table S6: Chlorophyll (Chl), HCN (dhurrin) and nitrate (NO<sub>3</sub>) concentration, and total nitrogen allocation in the leaf tissue of well-watered and drought stressed *Sorghum bicolor*  and forage sorghum HyA. Mean ± standard error and significance of a two-way ANOVA are presented for each variety (V) and treatment (T).



WW – well watered; WL – water limited. \*P< 0.05; \*\*P< 0.01; \*\*\*P< 0.001; ns, not significant