Using Bayesian Growth Models to Predict Grape Yield

Rory Ellis

Supervisors: Daniel Gerhard, Elena Moltchanova

Background

- Seasonal differences in vine yield need to be managed to ensure appropriate fruit composition at harvest.
- Weather conditions at flowering can cause knock-on events later in the growing season.
- Therefore early indications of yield are important in knowing what management practices must be undertaken during to give optimum yield.

Aims

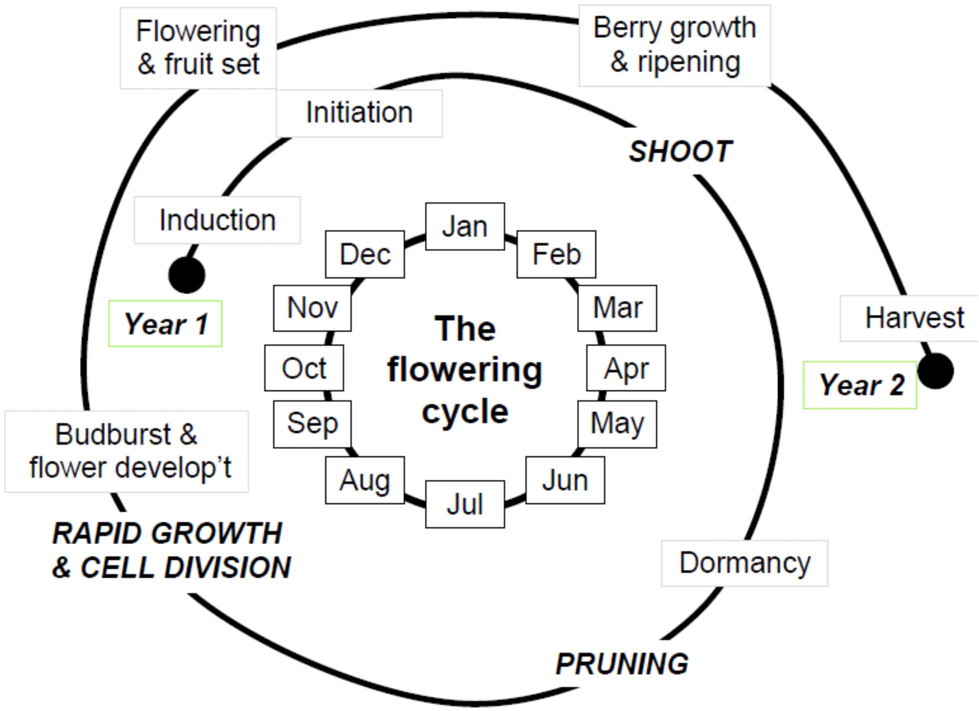
- Develop a tool which can assist New Zealand vineyard owners in performing early predictions of grape yield.
- Improve upon current yield estimation practices in the industry.

Introduction

- Understanding Grape Growth
- Bayesian Model Analysis
  - Derivation of Priors
  - Studies involving value of added information and vague vs. informed priors
- Results
- Conclusions
- Future Work

Understanding how grapes grow

- Grapevines are a biennial plant (two year growth cycles).
- Generally, grapevine yield develops over a 15-month period.
- During the growing season up to harvest, the grape berries venture through many growth phases:
  - An initial period where flowers change to fruits,
  - A lag phase, which leads up to berry ripening (*véraison*),
  - A second growth phase where sugars and water start to accumulate in the berry.

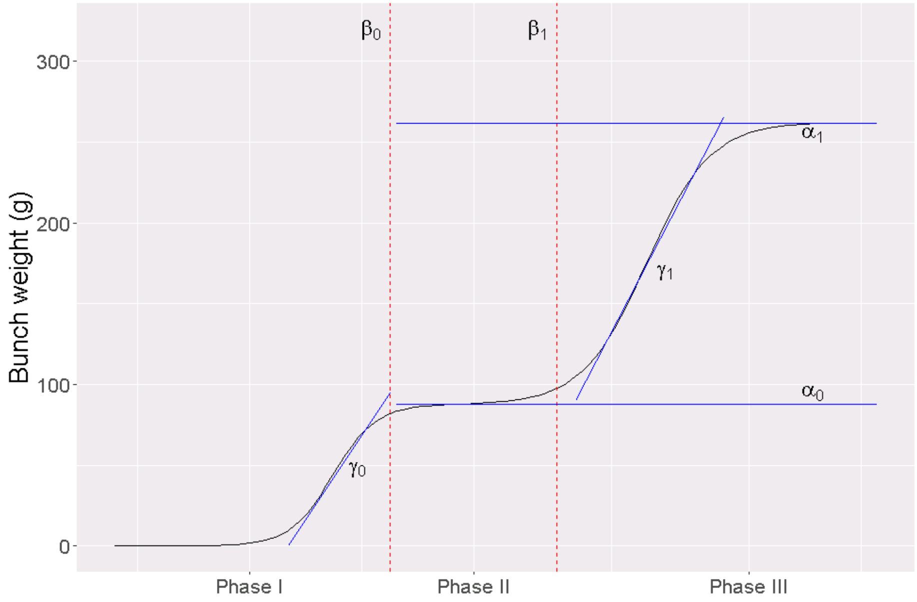


Grape Growth Characteristics

- Early work in understanding the phenology of grapes indicated a double sigmoidal growth pattern (Coombe, 1976).

$$\text{Yield}_i = \frac{\alpha_0}{1+e^{-\gamma_0(t_i-\beta_0)}} + \frac{\alpha_1}{1+e^{-\gamma_1(t_i-\beta_1)}} + \varepsilon_i$$

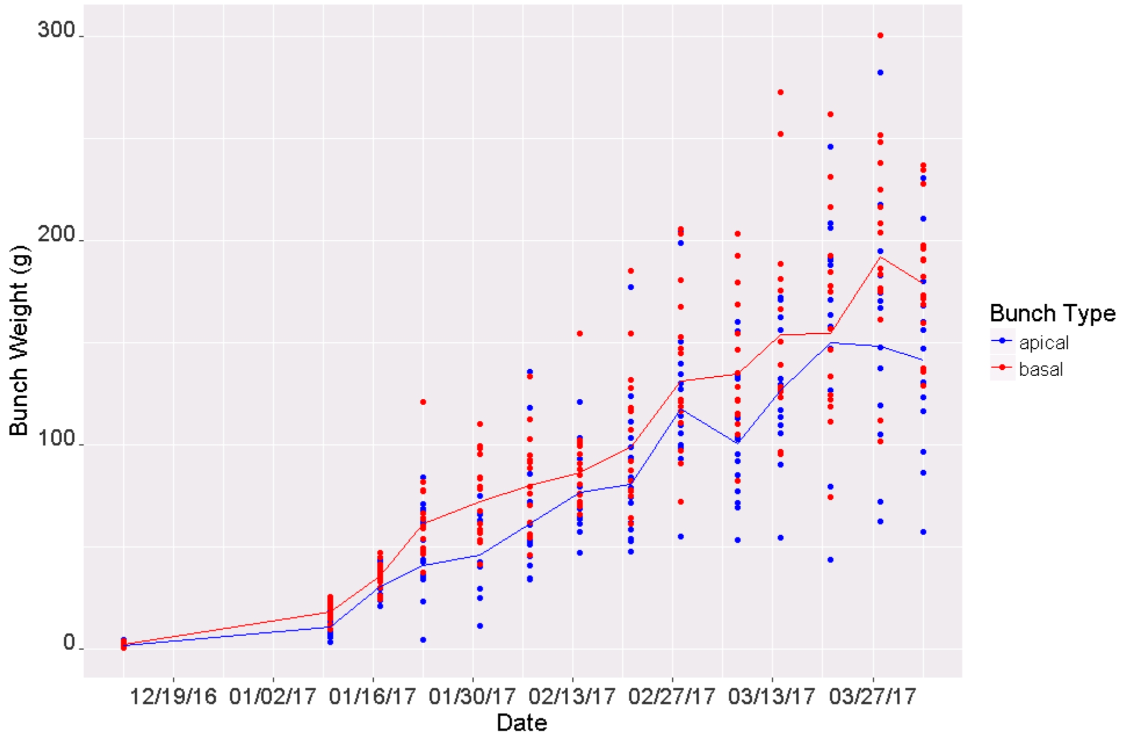
- α_0 and α_1 are the asymptote parameters
- γ_0 and γ_1 are the slope coefficients
- β_0 and β_1 identify the location of the inflection points
- $\varepsilon_i \sim N(0, \sigma^2)$ is the error term for the model.



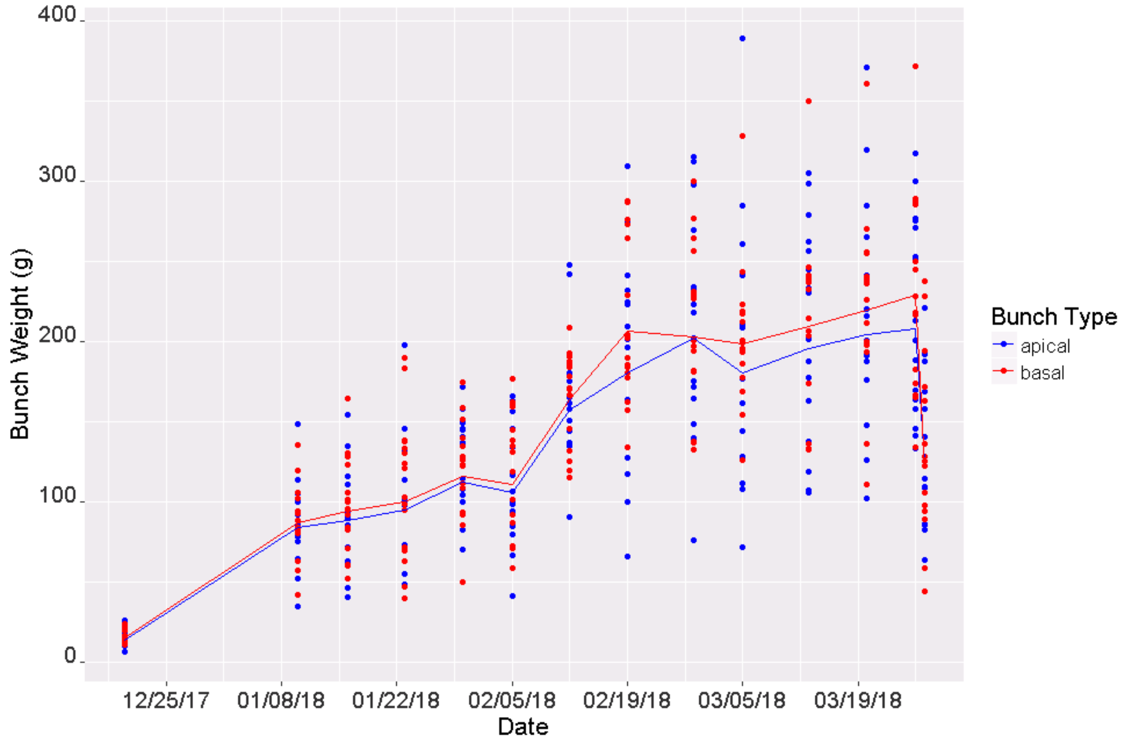
Bunch Weight Data

- 30 grape bunches in 2016/2017, and 32 grape bunches in 2017/2018, were destructively sampled. Weight measurements taken, at 14 time points throughout the respective growing seasons.
- Half of the bunches were taken from apical shoots, and half from basal shoots.

2016/2017 Growing Season



2017/2018 Growing Season



Bayesian Model Analysis

- One aim of this work is to determine the Value of Information (VOI) for this area of analysis.
- Double Sigmoidal Model is re-fitted iteratively with the inclusion of each new day's bunch weight.
  - Do the final yield estimates improve significantly upon doing this?

- The other aim involves determining the impact of incorporating *historical* data into the modelling procedure.
  - Will having yield data from previous years improve the estimates for the current growing season?

Derivation of Priors

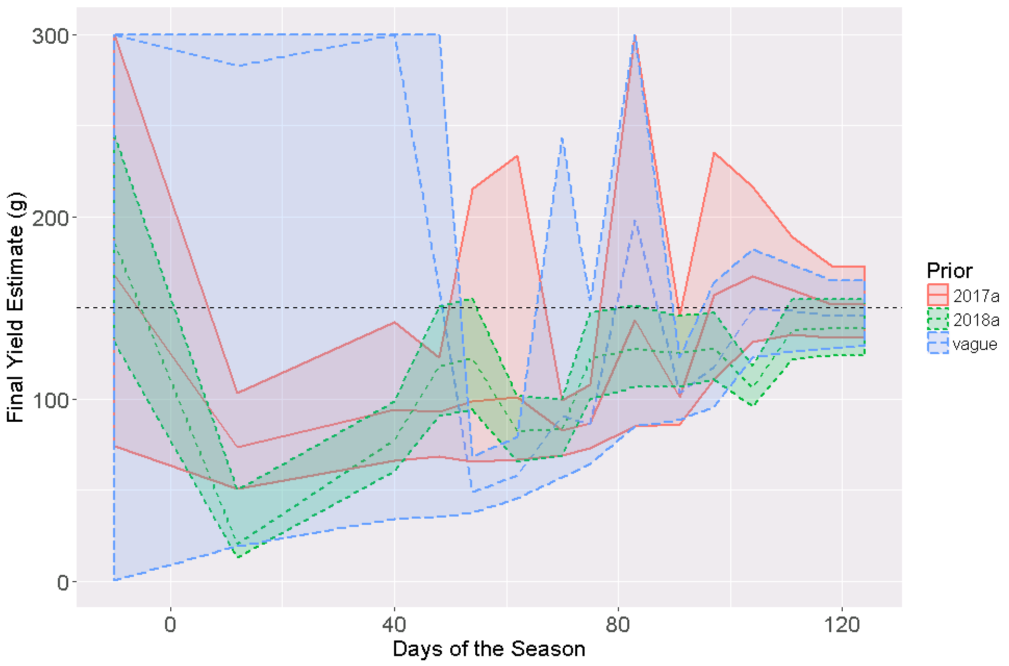
- Firstly, a set of weakly informative priors were found by assessing the shape of the grape growth curve *(on a log scale)*, and allowing for small precision (high variance) in each of the parameters.

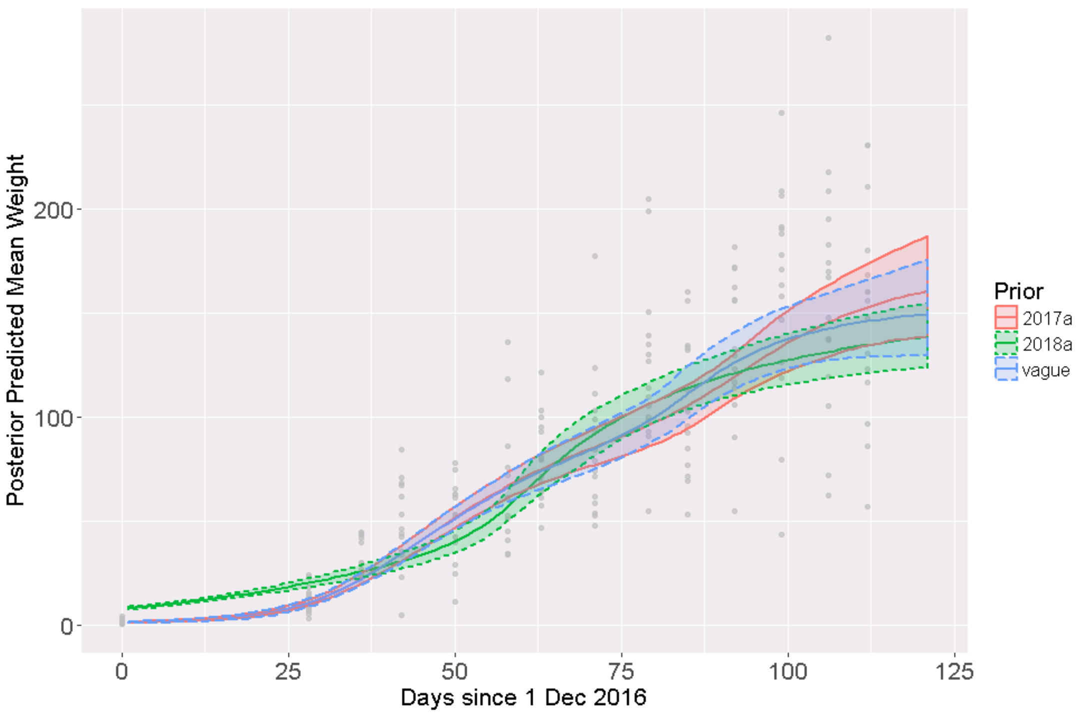
Coefficient	Prior
α_0	N(4.09, 0.11)
$\Delta\alpha$	TN(0.69, 4, 0)
β_0	N(35, 0.02)
$\Delta\beta$	TN(49, 0.11, 0)
γ_0	TN(0.3,44.44, 0)
γ_1	TN(0.3, 44.44, 0)
τ	Gamma(4,1)

N = Normal Distribution
TN = Truncated Normal Distribution

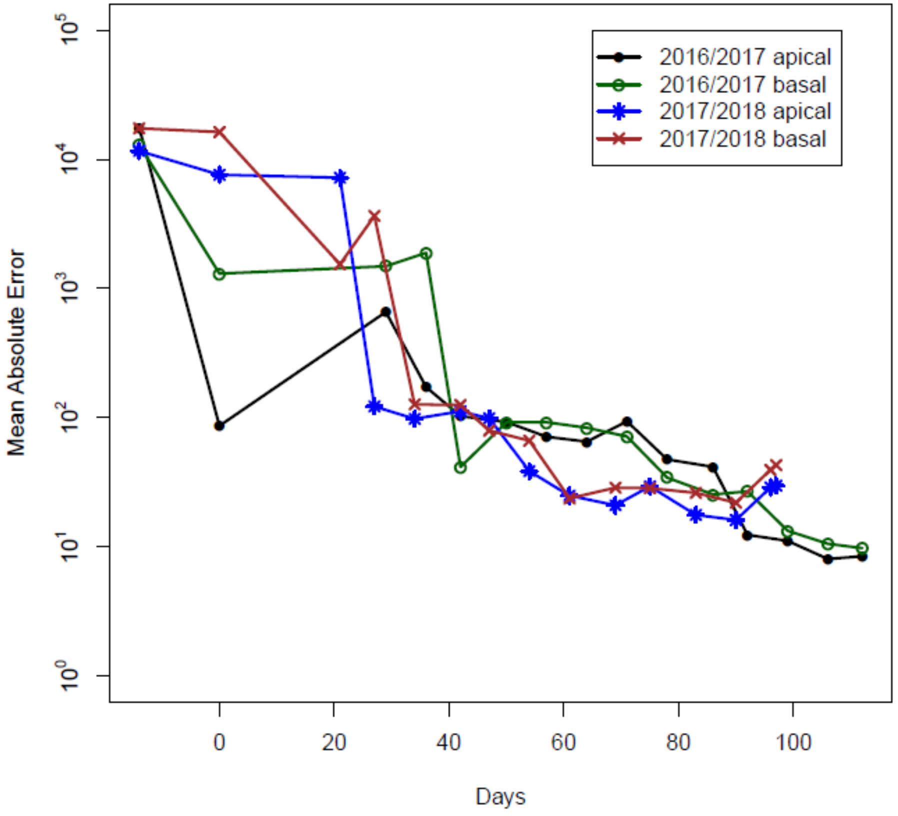
- Two other sets of priors were obtained via parametric approximation of the posterior distributions from analyzing the 2016/2017 and 2017/2018 apical bunches respectively.

Comparing yield estimates with different priors





Mean Absolute Error measurements comparing final yield estimates

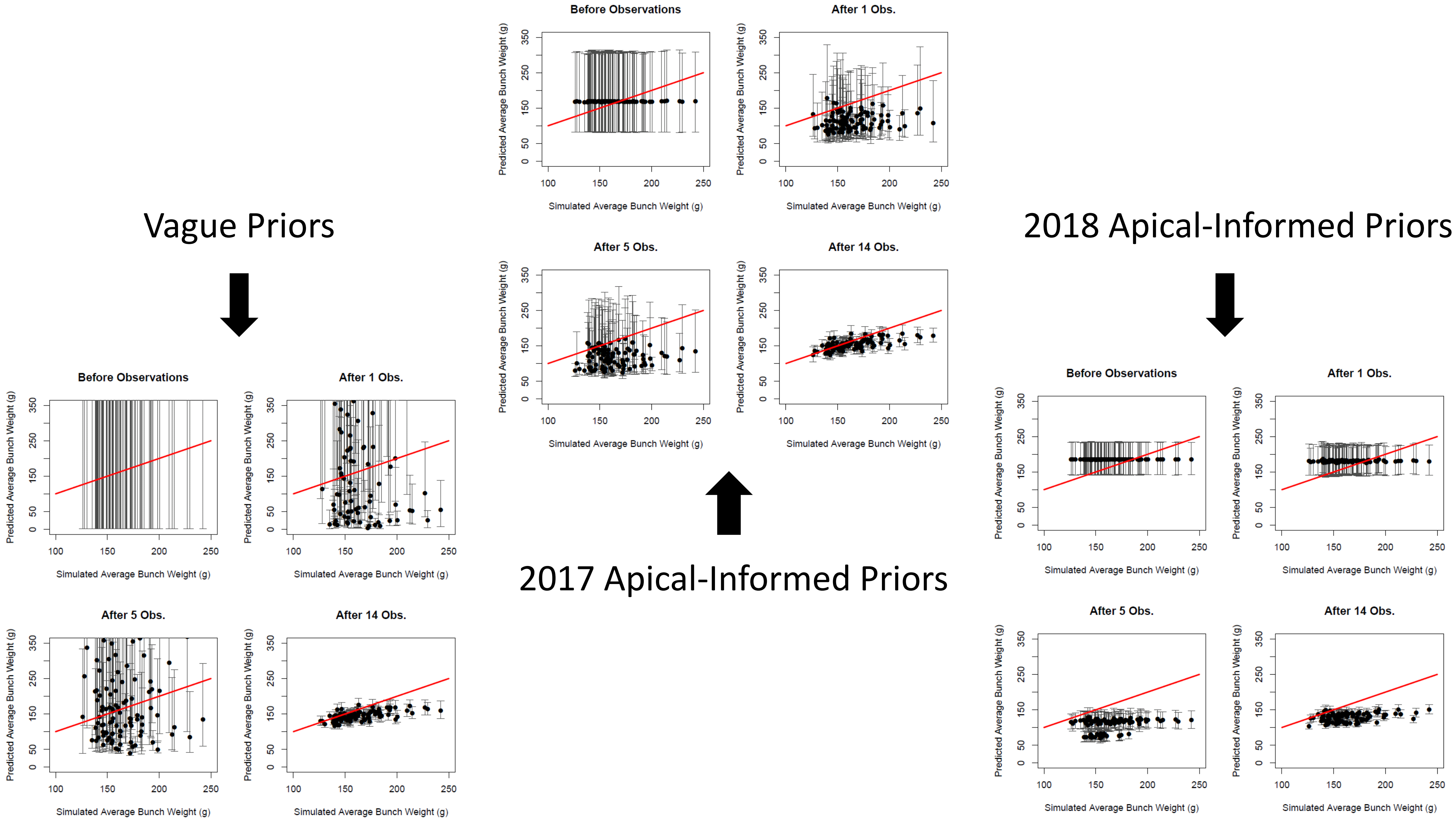


- The vague priors are fitted to the Bayesian Model.
- MAE measures comparing final yield estimates with the actual yield for the respective groups were found iteratively.
- Value of Information is obvious here.

Simulation Studies

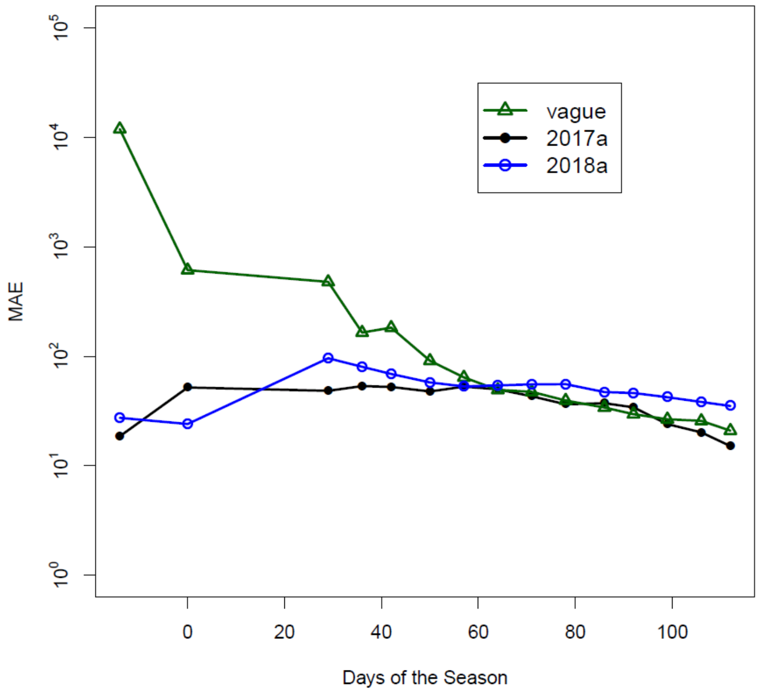
- Having data of the quality we have is hard to come by:
  - Vineyard owners typically do not conduct weekly destructive measurements of their bunches for a variety of reasons.

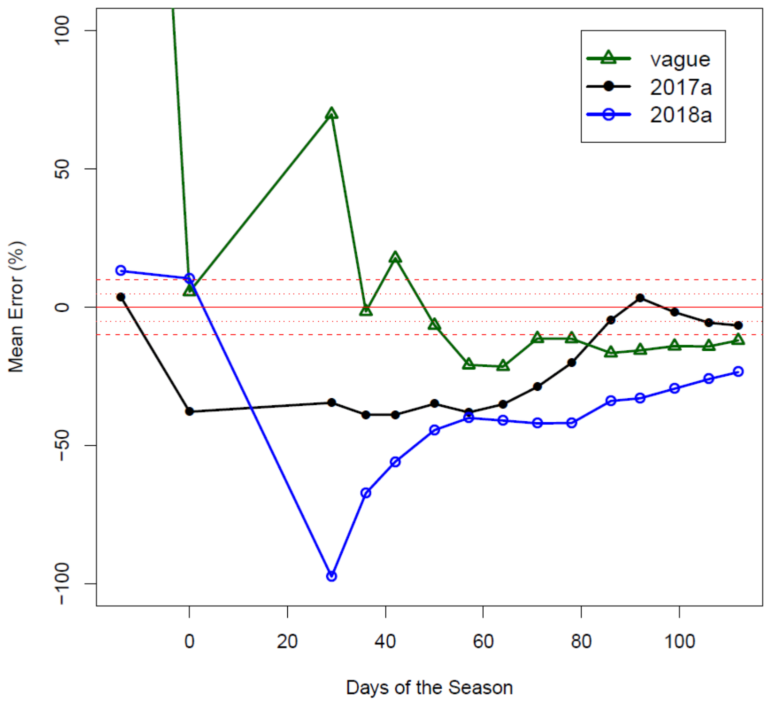
- 100 data sets based on the parameters derived from the 2016/2017 apical bunch weight data were simulated
- Priors derived previously then fit to the Bayesian model in each case, and MAE, MPE measures, alongside finding the 95% credible intervals



Vague Priors
Before Observations
After 1 Obs.
After 5 Obs.
After 14 Obs.
Predicted Average Bunch Weight (g)
Simulated Average Bunch Weight (g)
2017 Apical-Informed Priors
Before Observations
After 1 Obs.
After 5 Obs.
After 14 Obs.
Predicted Average Bunch Weight (g)
Simulated Average Bunch Weight (g)
2018 Apical-Informed Priors
Before Observations
After 1 Obs.
After 5 Obs.
After 14 Obs.
Predicted Average Bunch Weight (g)
Simulated Average Bunch Weight (g)

Mean Absolute Error and Mean Percentage Error results.





Conclusions

- In these studies, the Bayesian Model is sensitive to prior assumptions.
- Having a non-informative (vague) prior may be more beneficial in producing final yield estimates, than having informed priors based on one unusual year.
- Evident trade-off between early final yield prediction vs. accurate final yield prediction.
- A Bayesian framework is useful in this context, due to its ability to update model estimates as new data comes in.
  - This is important due to the dynamical nature of grape growth.

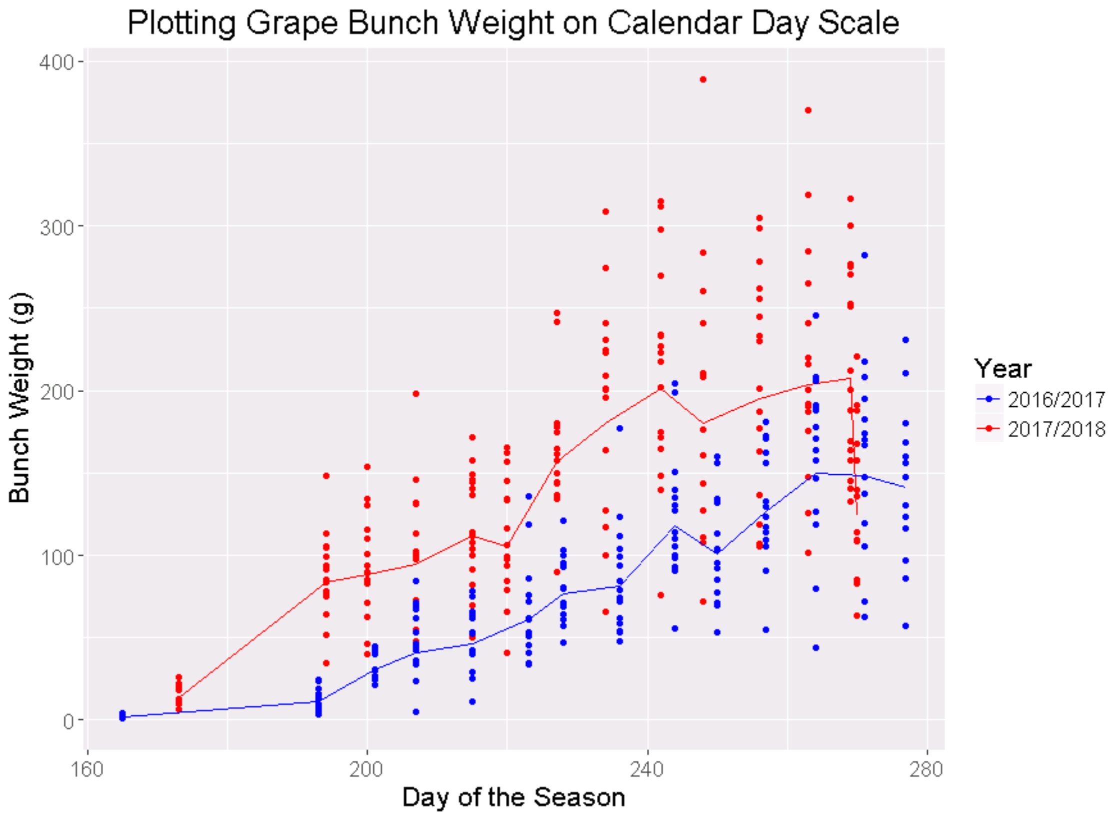
Future Work

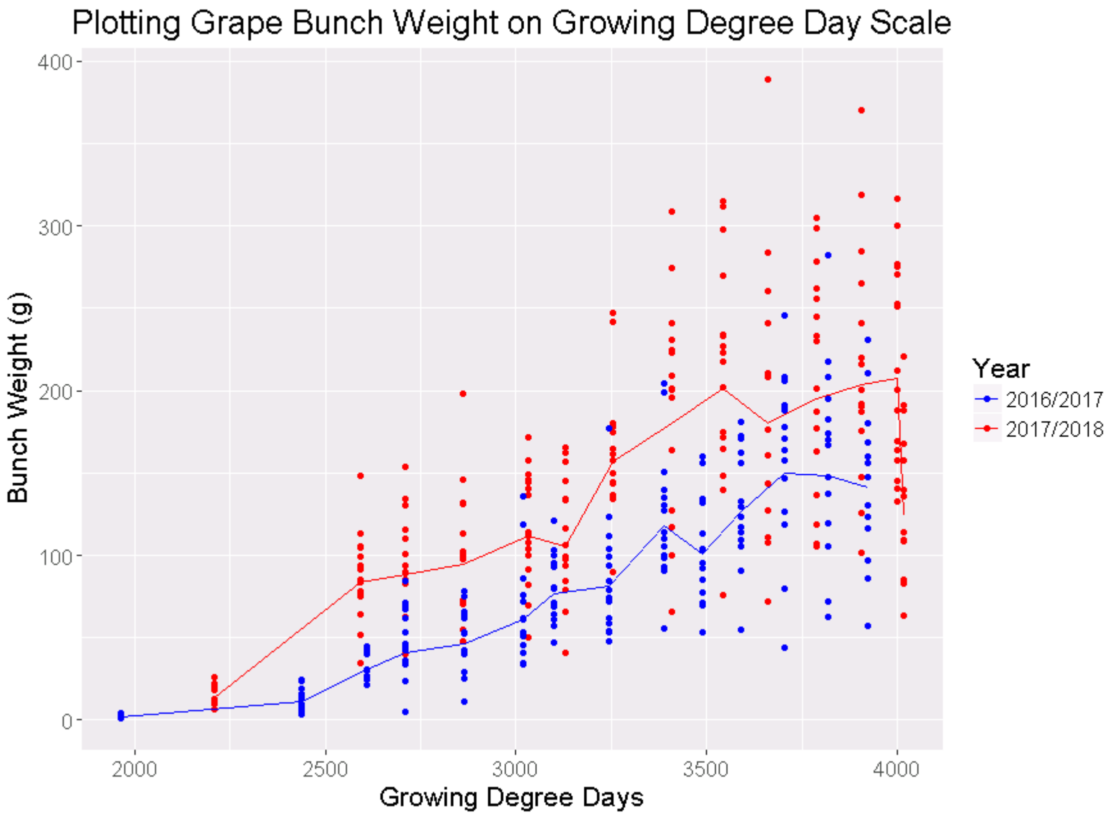
- Publication of this work.
- Modelling the bunch weight data on a temperature scale (Growing Degree Days).
- Analyzing climatic impacts in Bayesian Modelling procedure.
- Implementing MCMC methods (Metropolis-Hastings Algorithm) to estimate the Bayesian Model.
- Explore other nonlinear model specifications

Climatic Impacts

- There are two ways of considering this:
  - Firstly, can the bunch weight data be modelled on a scale relating to temperature, instead of calendar days?
  - Secondly, would factors like temperature, rainfall, or solar radiation impact the parameters of the Bayesian model?
- Growing Degree Days are found by summing the average temperatures over the sequence of days up to the day of measurement.
  - Starting point is July 1 in the Southern Hemisphere. e.g. 2016/2017 growing season begins on July 1 2016.

Comparison of Calendar Day and Growing Degree Day Scale





Different Specifications of the Double Sigmoidal Curve

- The current double logistic model fits well to the 2017/18 growing season data in particular.
- One particular issue may be asymmetry in the grape growth during the growing season.
- There are other model specifications which combat this issue.

Different Specifications of the Double Sigmoidal Curve:

- **5-parameter Logistic :**
  - $y = \frac{\alpha_0}{1+e^{-\gamma_0(t-\beta_0)^{e_0}}} + \frac{\alpha_1}{1+e^{-\gamma_1(t-\beta_1)^{e_1}}} + \varepsilon_i$

e_0, e_1 = asymmetry parameters

- **Richards (Richards 1959):**
  - $y = \frac{\alpha_0}{(1+k_0 e^{-\gamma_0(t-t_{m0})})^{1/k_0}} + \frac{\alpha_1}{(1+k_1 e^{-\gamma_1(t-t_{m1})})^{1/k_1}} + \varepsilon_i$

k_0, k_1 fix point of inflection
t_{m0}, t_{m1} = time of maximum growth

- **Gompertz (Gompertz 1825):**
  - $y = \alpha_0 e^{e^{-\gamma_0(t-\beta_0)}} + \alpha_1 e^{e^{-\gamma_1(t-\beta_1)}}$
- **Weibull (Weibull 1951):**
  - $y = \alpha_0\left(1-e^{-\left(t/\beta_0\right)^{c_0}}\right) + \alpha_1\left(1-e^{-\left(t/\beta_1\right)^{c_1}}\right)$

c_0, c_1 = shape parameters

Acknowledgements

Funded by: MBIE LVLX1601 and NZ Wine

MINISTRY OF BUSINESS, INNOVATION & EMPLOYMENT
HIKINA WHAKATUTUKI

NEW ZEALAND WINE
PURE DISCOVERY

Project initiated and co-ordinated by:

LINCOLN AGRITECH LTD

Further project partners and contributors to this presentation:

Lincoln University
Te Whare Wānaka o Aoraki
AOTEAROA • NEW ZEALAND
New Zealand's specialist land-based university

Plant & Food RESEARCH
RANGAHAU AHUMĀRA KAI

UC
UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

CSIRO

Acknowledgements

- Mike Trought and LinLin Yang (Plant and Food Research).
- Armin Werner (Lincoln Agritech).
- Amber Parker (Lincoln University).

References:

- Coombe, B. (1976), ‘The development of fleshy fruits’, Annual Review of Plant Physiology 27(1), 207–228.
- Gompertz, B. (1825), ‘Xxiv. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. in a letter to Francis Baily, Esq. frs &c’, Philosophical transactions of the Royal Society of London 115, 513–583.
- Richards, F. (1959), ‘A flexible growth function for empirical use’, Journal of experimental Botany 10(2), 290–301.
- Weibull, W. (1951), ‘A statistical distribution function of wide applicability’, J Appl Mech 18, 293–297.