FINAL PROJECT REPORT **YEAR**: 3 of 3
WTFRC Project Number: AP-09-908

Project Title: Modeling Washington apple bloom phenology and fruit growth

PI:	Karen Lewis	**Co-PI (2):**	Tory Schmidt
Organization:	WSU Extension, Ephrata	**Organization:**	WTFRC
Telephone/email:	(509) 760-2263 kmlewis@wsu.edu	**Telephone/email:**	(509) 665-8271 tory@treefruitresearch.com
Address:	P. O. Box 37	**Address:**	1719 Springwater Ave.
Address 2:	Courthouse	**Address 2:**	
City:	Ephrata	**City:**	Wenatchee
State/Zip:	WA 98823	**State/Zip:**	WA 98801

Co-PI(3): Nairanjana Dasgupta
Organization: WSU Dept. of Statistics
Telephone/email: (509) 335-8645
dasgupta@wsu.edu
Address: Neill Hall #403
City: Pullman
State/Zip: WA 99164-3144

Cooperators: Tim Smith (WSU Extension, Chelan County), Gwen Hoheisel (WSU Extension, Prosser), Norman Suverly (WSU Extension, Okanogan), Tom Auvil (WTFRC), Ines Hanrahan (WTFRC), Felipe Castillo (WTFRC), Mark Bell (WTFRC), Ute Chambers (WSU-TFREC), Monte Shaffer (WSU, Pullman), Brandon Mulvaney (WTFRC), Udell Mendoza (WTFRC)

Total project funding request: **Year 1:** $4,180 **Year 2:** $7,938 **Year 3:** $5,690

Other funding Sources: None

WTFRC Collaborative expenses:

Item	**2009**	**2010**	**2011**
Stemilt RCA room rental			
Crew labor[1]	5,000	5,000	7,000
Shipping			
Supplies			
Travel[2]	1,800	1,800	2,400
Miscellaneous			
Total	$6,800	$6,800	$9,400

[1] Labor calculated as 2 persons at $16.00/hr working 12 hrs per week for 13 weeks during the growth season.
[2] In-state travel to research plots.
NOTE: 2011 budget increased from $6800 to reflect additional sites handled by WTFRC staff

Budget 1

Organization Name: WSU Extension **Contract Administrator:** M.L. Bricker
Telephone: (509) 335-7667 **Email address: mdesros@wsu.edu**

Item	2009	2010	2011
Salaries[1]		2,941	3,059
Benefits		847	881
Wages[2]	1,000	1,000	1,000
Benefits	180	150	150
Equipment			
Supplies			
Travel[3]	3,000	3,000	600
Total	$4,180	$7,938	$5,690

1 Salary (benefits at 28.8%) for Nairanjana Dasgupta
2 Wages (benefits at 15%) for part-time help in Wenatchee for bloom observations.
3 Cooperator in-state travel for bloom observations (5 persons at $600 each)
NOTE: 2011 budget decreased from $8090 to reflect reduced participation from Extension staff (Olmstead, Suverly, Lewis, Hoheisel)

Objectives:

1. Develop functional models for apple bloom development from bud break to petal fall for three cultivars: ‘Red Delicious’ (standard with historic data), ‘Cripps Pink’ (early bloomer), and ‘Gala’ (mid-late bloomer).
2. Develop fruit growth models for the same three cultivars from petal fall until harvest.
3. Incorporate models into WSU DAS system.

Significant Achievements:

- Bloom phenology observations successfully recorded at 11 location nodes throughout Central Washington, including 11 Red Delicious, 11 Gala, and 9 Cripps Pink blocks (Table 1)

- Fruit diameter measured throughout growing season at 11 Red Delicious, 10 Gala, and 9 Cripps Pink blocks; fruit length also recorded in Red Delicious blocks (Table 1)

- Growth data is being modeled using a 3 parameter Richard’s curve using non-linear mixed modeling approach accounting for the repeated observation on the apples (Table 2); the model predicts data well, but seasonal differences (between 2010 and 2011) indicate further refinement would be beneficial; analysis of 2011 data, including temperature data, is ongoing

- Bloom data is modeling using ordinal logit model with each bloom phase as a response. Here we model the phase of growth depending on the days after ½ inch green phase; 2010 data shows promise for this model (Tables 4-7); analysis of 2011,including temperature data, data is ongoing

- Accuracy of bloom phenology models were improved by factoring in degree days relative to two temperature thresholds, 32F and 50F (Tables 4-7)

- New proposal developed for beta-test evaluation and incorporation of models onto WSU AgWeatherNet (AWN) under leadership of Hoogenboom and Salazar

Project Timeline:

2009

- Project team assembled with WSU Extension and WTFRC internal staff
- Jim Olmstead (WSU) and Tory Schmidt (WTFRC) designated as co-PIs
- Field sites secured throughout Central Washington for data collection
- First season of bloom phenology and fruit growth data collected
- Olmstead left WSU; Schmidt assumed project leadership with help from Karen Lewis (WSU)
- Statistician for project (Nairanjana Dasgupta – WSU, Pullman) recruited

2010

- Data collection protocols revised based on recommendations from statisticians
- WTFRC internal staff assumed bloom observation & data collection at all but one site
- Second season of bloom phenology and fruit growth data collected
- Automated field cameras proved inadequate for bloom phenology observations
- Preliminary models built by statisticians

2011

- Third season of bloom phenology and fruit growth data collected
- WSU DAS deferred model validation & beta testing to WSU AWN
- Models updated/improved to incorporate new data, including temperature/degree days
- Hoogenboom and Salazar (WSU AWN) join team to develop new project proposal

Methods:

Bloom phenology: Team members from WSU Extension and WTFRC internal program observed and evaluated flagged apple blocks around the state (Table 1) at regular intervals from bud break until mean fruitlet size reached 20mm. Representative buds/clusters at chest level on the northwest side of trees of each cultivar were categorized by phenologic stage and digital pictures were taken of representative buds/flower/fruitlets. Based on input from WSU statisticians, observation intervals were shortened to 2-3 days (2-5 days in 2009) and sample size was increased to 30 buds for the 2010 and 2011 seasons (20 buds in 2009). Data were recorded on a tally sheet by each individual and eventually submitted to the WTFRC internal program for collation. Hobo data loggers were deployed at each site to record ambient temperatures throughout the season; weather data was dumped from the data loggers in June and collated for transfer to the statisticians.

Fruit growth: After June drop and hand-thinning, 50 surviving fruit were tagged in the same blocks used for the bloom phenology observations. All fruit were measured by WTFRC staff with for diameter and Red Delicious was additionally measured for length as an indicator of fruit type at weekly intervals until the blocks were harvested in the fall. Weather data was gathered from Hobo data loggers after harvest at each site, collated, and submitted to the statisticians for analysis. As with bloom phenology, fruit growth protocols (sample size and intervals) were modified for 2010 and 2011 based on recommendations from statisticians.

Table 1. Roster of sites utilized for apple bloom phenology observations and fruit growth measurements. 2009-2011. (RD = Red Delicious, CP = Cripps Pink, G = Gala)

LOCATION	GROWER	CVs	ELEV (ft)	STAFF	FRUIT GROWTH
Omak[1]	Root	RD, G	1250	Suverly/Crew	RD only
S Shore Chelan	Easley	CP	1120	Auvil/Crew	Y
	Sunshine	RD, G	1450	Auvil/Crew	Y
Brays Landing	Podlich	RD, CP, G	900	Auvil/Crew	Y
S Orondo	C & O Nursery	RD, CP, G	755	Crew	Y
E Wenatchee	Gausman	RD, CP	910	Esteban	Y
	Witte	G	1025	Esteban	Y
Rock Island	WSU-TFREC	RD	910	Crew	Y
	WSU-TFREC	G	880	Crew	Y
	Zirkle CRO	CP	775	Crew	Y
Royal Slope	Delay	CP	1095	Lewis/Crew	Y

	Delay	RD, G	1055	Lewis/Crew	Y
Naches	Rowe	RD, G	1580	Crew	Y
Parker	Brandt	RD, CP, G	879	Crew	Y
Sawyer	WTFRC Rootstock	G	870	Crew	Y
	Badgely	RD	870	Crew	Y
	Weippert	CP	870	Crew	Y
Prosser	Ballard	RD, CP, G	681	Hoheisel/Crew	Y

[1] Omak site discontinued in 2011 after departure of Suverly from project

Results & Discussion:

Data collection. Despite the losses or revised roles of team members over the course of the project, we successfully collected 3 seasons of solid data to help construct these models. Because we had to initiate the project without the counsel of an experienced statistician or modeler, our 2009 data collection efforts reflected some inefficiencies that were corrected for the 2010 and 2011 seasons. Based on recommendations from Dasgupta and her students, the following changes were adopted:

- Shorter and more regular sampling intervals
- Increased sample size for bloom observations
- Decreased sample size for fruit measurements
- Standardized data collection protocols

While the 2009 data has value, it is not as complete or robust as 2010 and 2011 data in terms of statistical strength.

Hobo data loggers were again deployed at all nearly all sites to record ambient temperatures in the immediate microclimate of the sampled trees; most sites were selected due to their proximity to AWN stations (usually within a mile), and models using temperatures from both systems could be evaluated for the best statistical fit. Potential discrepancies between temperatures recorded by AWN and individual data loggers could have many explanations, but may be instructive regarding broader extrapolation of weather readings from either system.

In an effort to explore options for reducing time commitments for our field personnel, we tested autonomous digital cameras designed for monitoring big game trails to assess their utility for making routine observation of bloom development in 2010. Unfortunately, the effective focal ranges for these cameras are 5+ feet; images taken of branches inside that range proved too blurry to be useful and branches that were in adequate focus were too far from the camera to discern details of individual buds or flowers.

Modeling. *Fruit growth:* Our pilot data from 2009 and literature survey indicated we use a non-linear regression model to model the growth pattern for the different locations for the 2009 data. We used the Richards's curve formulation (model given below):

$$Y_{ji} = \frac{\beta_i + uj}{(1 + e^{\delta_i(X_{ji} - \tau_i)}}$$

Where: Y_{ji} represents the growth for apple *j* at location *i*,
X_{ji} denotes the time in Julian Days,
β represents the maximal growth reached by the fruit,
δ represents the growth rate

τ represents time when maximum growth occurred.

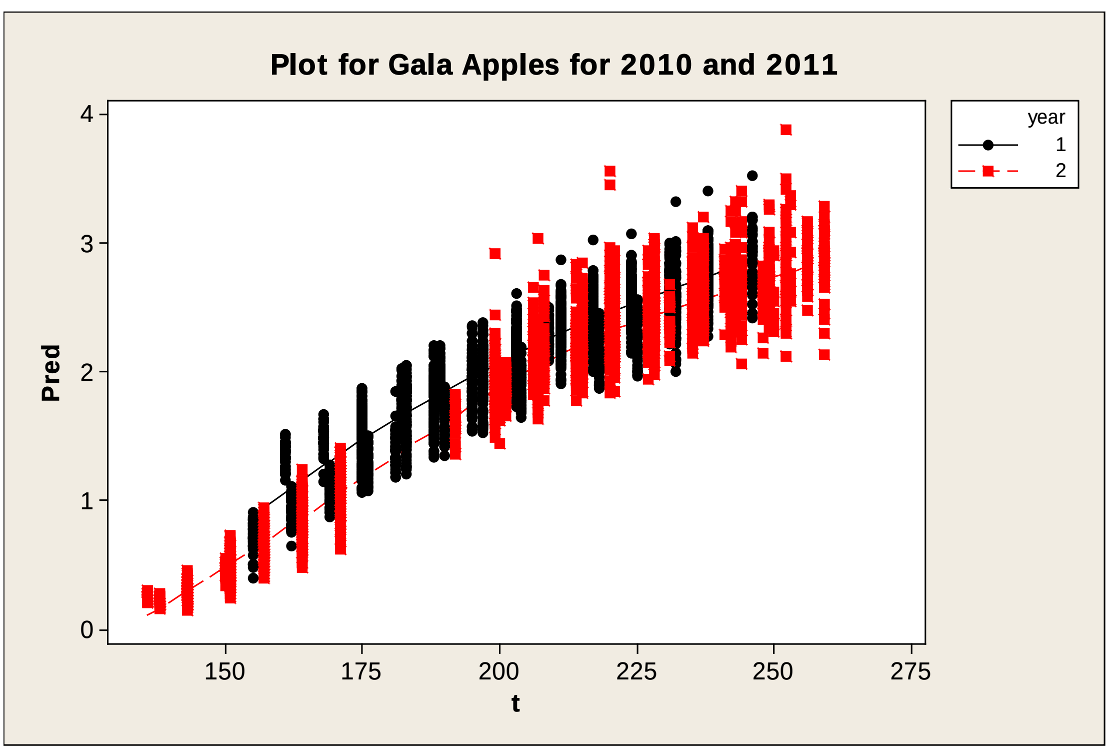
As these parameters all have physical meaning in the context of apples we decided to use this model. Our initial finding showed significant differences across sites. To compare across sites we used a technique that we developed (Many-to-one comparisons for Apple Growth, Dasgupta and Shaffer, submitted to *Journal of Applied Statistics*).

The 2009 growth data was used mainly to get an idea about the scheme for data collection and was used mainly as the pilot study for this project. However data for 2010 and 2011 were used actively for the modeling purpose. In both years we ran the Richard's curve model given earlier. One reason for this choice was the easy interpretability of the parameters as mentioned in the modeling section.

Table 2. Gala data for 2010 and 2011 over all locations

parameter	Estimates from 2010	Estimates from 2011	Estimates combining the two years
β (maximal growth)	3.16	2.68	2.73
δ	179.34	182.38	177.25
τ	.03	.047	.045
Variation due to apple	.093	.1355	.1105
Random variation	.0088	.0034	.06

Table 2 indicates that there were differences in the parameters across both years. In 2010 the estimated maximal size of the apple was predicted bigger than that in 2011. Combining the two years gives us less total variability, but it may not be prudent to combine the data from two different years. The graph below shows the predicted values for Gala for 2010 and 2011 (as year 1 and 2). It can be seen that the Gala apples in 2010 were larger than their 2011 counterparts.



Bloom phenology: We fitted an ordinal logistic regression with the stage as our response variable and put in Julian date, various temperature readings, and location as explanatory variables in the model. This analysis uses a proportional odds model with cumulative logit link and is defined as

$$\ln\left(\frac{P(Y \le i)}{P(Y > i)}\right) = \alpha_i + \beta X \qquad i = 1,2,\dots,k-1$$

where i represents the first 7 categories (bloom stages) which are compared to the 8th stage. The intercept α_i is different for each stage, however the model assumes common slopes between each of the k-1 = 7 regression lines, meaning the cumulative logit comparing the first 7 stages to the 8th stage increases at the same rate across all predictor variables.

We had at first wanted to have a model that incorporated cultivar as an explanatory variable, but we realized that there was interaction between cultivars and Julian date and this would further complicate our model. In Table 3, we see that the "most likely" stage of the cultivar based on Julian Date is different across the cultivars. We computed the "most likely" stage using our observed data over all the locations.

Table 3. "Most likely" stage based on Julian date that was used in the exploratory phase

Cultivar	Red Delicious	Gala	Cripps Pink
Stage			
Green Tip	72-79	72-79	72-74
½ Inch Green	81-88	81-96*	75-85*
Tight Cluster	89-96	97-104*	86-90*
First Pink	97-105*	105-106*	92-97*
Full Pink	106-107*	107-110	98-104
First Bloom	108-109	111-113	105-107
Full Bloom	109-113	111-113	108-110
Pedal Fall	>113	>113	>110

From the table it is clear that between cultivars there is large variation for the days they are in a particular stage, though some of this variability may be due to locational differences. Because growers would be interested in predicting the stages of the cultivars separately, it was decided to obtain a model for each different cultivar. Hence we ran our model for different combinations of temperature, location and Julian date to obtain the simplest model with the best predictions. The purpose was to see if good predictions could be obtained without the location variable as that would make the models more general and more usable by the growers who might not be in a specific location used in our study.

The best models determined by the goodness of fit tests were compared to one another in the context of other measures of model fit. AIC and BIC, which award a model for fitting well but penalize for over overparameterization (BIC penalizes more), R^2 which measures the variability in the data explained by the model, and concordance percent, which is the percent of the stages that were correctly predicted by the logistic model. The results for the best models for the three cultivars are given in tables below:

Table 4. Bloom phenology model performance for Red Delicious

Red Delicious					
Model	Goodness of fit	AIC	BIC	R^2	Concordance (%)
Location, >50, >32, Day	9.317577982	6339.943	6448.135	0.9393	97.2
Location, >50, >32	9.458133537	6339.923	6441.752	0.9393	97.2
Location, >50, Day	9.52692322	6344.385	6446.213	0.9392	97.1
Location, >50	9.52692322	6342.385	6437.849	0.9392	97.1
Day, >50, >32	12.86715709	7234.05	7297.693	0.9207	96
Day, >50	12.59888102	7674.874	7732.126	0.9101	95.7

To see the effect of location we present the following table:

Table 5. Red Delicious predicted stages using "Location, >50, >32" model on the same days across locations

Julian Day	Chelan	Bell	Konnowac Pass	Naches	Olmstead	Smith	Sunrise	Omak
99	B	C	E	B	D	C	C	B
106	C	D	F	D	E	E	D	C
116	E	G	H	F	H	H	G	F

This result is mirrored with the other cultivars with only slight differences between model fits:

Table 6. Bloom phenology model performance for Gala

Gala					
Model	Goodness of fit	AIC	BIC	R^2	Concordance (%)
Location, >50, >32, Day	6.98234159	4892.717	5001.008	0.9567	98.1
Location, >50, >32	6.990776934	4903.697	5005.618	0.9565	98.1
Location, >50, Day	7.605574514	4912.656	5014.578	0.9564	98.1
Location, >50	7.66044118	4914.132	5009.684	0.9563	98.1
Day, >50, >32	12.09768221	6720.874	6784.575	0.9226	96.5
Day, >50	14.78781519	6871.594	6928.925	0.919	96.2

Table 7. Bloom phenology model performance for Cripps Pink

Cripps Pink					
Model	Goodness of fit	AIC	BIC	R^2	Concordance (%)
Location, >50, >32, Day	2.735705285	3749.901	3841.114	0.9533	98
Location, >50, >32	3.030880182	3756.072	3841.204	0.9531	97.9
Location, >50, Day	3.344702934	3854.65	3939.782	0.951	97.8
Location, Day, <32	3.452095538	4271.446	4456.578	0.9388	97.1
Location, >50	3.195721452	3878.671	3978.671	0.9505	97.8
Day, >50, >32	7.802714615	5408.775	5469.584	0.9072	95.4
Day, >50	10.85301005	5911.777	5966.504	0.8878	94.7

The overall finding from the 2010 bloom data is that bloom is best predicted using location, Julian date, temperature over 50 and temperature below 32 as predictors.

Beta testing/web integration. As stated in Objective 3, our original hope was to develop preliminary models that could be evaluated on the WSU Decision Aids System (DAS). While the DAS team has been supportive of developing these two models, they expressed some reservations regarding their lack of experience with validation of horticultural models and suggested that working with Gerrit Hoogenboom and the AgWeatherNet (AWN) team might better serve our needs. In fact, Hoogenboom would have been an obvious choice as a co-PI on this project if he were available to us in 2009. Nonetheless, initial discussions with Hoogenboom and Melba Salazar (AWN) have been productive and we have developed a new proposal for the extension of this project to strengthen and evaluate the models using AWN as a platform for beta testing.

EXECUTIVE SUMMARY

This project was initiated in 2009 as a joint effort between WSU Cooperative Extension and the WTFRC internal program to develop a phenologic model which could help industry improve resource management decisions during spring with better prediction of apple bloom development based on weather forecasts. Three cultivars were selected for study: 1. Cripps Pink (early bloomer) 2. Gala (mid-late bloomer) 3. Red Delicious (historical standard). Observation sites for these three cultivars were also used to collect fruit growth data from June drop until harvest to develop models which can predict the growth curves and ultimate harvest size of these fruit.

Data were collected for both models over three seasons (2009-2011) at 11 nodes across Central Washington. Weather data was recorded by data loggers at most sites, as well as AWN stations near all sites. These data were collated, formatted, and submitted to statisticians at WSU (Pullman) for development of the respective models.

Preliminary bloom phenology and fruit growth models show promise, but require more data and further statistical refinement to better incorporate weather data and explore variability between cultivars, sites, and years. Data analyses indicate that incorporation of degree days relative to two temperature thresholds (32F and 50F) improve the accuracy of model predictions.

After consultation with DAS staff at WSU-TFREC, it was decided to pursue beta-testing and online model integration with AWN. A new project proposal has been submitted to WTFRC detailing this proposed work.