Development of a Prototype Bayesian Network Model Representing the Relationship between Fresh Market Yield and Some Agroclimatic Variables Known to Influence Storage Root Initiation in Sweetpotato

Arthur Villordon[1]
LSU AgCenter Sweet Potato Research Station, 130 Sweet Potato Road, Chase, LA 71324

Julio Solis and Don LaBonte
LSU AgCenter School of Plant, Environmental, and Soil Sciences, 137 J.C. Miller Hall, Baton Rouge, LA 70803

Christopher Clark
LSU AgCenter Department of Plant Pathology and Crop Physiology, Baton Rouge, LA 70803

Additional index words. Ipomoea batatas, storage root initiation, anomalous cambium, growing degree-days, phenology, decision support system

Abstract. A prototype phenology-driven Bayesian belief network (BBN) model, named BxNET, was developed to represent the relationship between fresh market yield (U.S. #1 grade) and agroclimatic variables known to influence the critical storage root initiation stages in 'Beauregard' sweetpotato. This data-driven model was developed from experimental data collected over 3 years of field trials in which management variables were kept as uniform as possible. The BBN was developed assuming that soil moisture measured at the 15-cm depth was not a limiting variable during the first 20 days after transplanting, during which the onset of storage root initiation determined the majority of storage root yield at harvest. The absence of influence from weeds, disease, insect pests, and chemical injury was also assumed. Accuracy of the fully parameterized working prototype was estimated through leave-one-out cross-validation (14% error rate), validation on an independent test data set (20% error rate), and area under the receiving operator characteristic curve (0.59) analysis. As a result of its empirical nature, BxNET is only applicable to the cultivar, location, and the limited set of environmental (air temperature, soil temperature, relative humidity, solar radiation) and management variables as defined in the 3-year study. This beta-level model can serve as a foundation for the development of a final working model through further testing and validation. Additional validation data may require revision of the current model structure and conditional probabilities. These validation studies will also allow the model to be used in other locations. BxNET can be expanded to include other causal variables such as weed incidence, disease presence, insects, and chemical injury. Such an expansion can lead to the development of a model-based decision support system for sweetpotato production. Such a system can help model alternative management scenarios and determine the most reasonable management interventions to achieve optimum yield outcomes under different agroclimatic conditions.

Singh et al. (1992) predicted that computer-based crop growth models were going to increase for root crops, including sweetpotatoes. To our knowledge, only two process-based models have been published thus far (Mithra and Somasundaram, 2008; Somasundaram and Mithra, 2008). Such models represent an important step in further understanding the complex interactive nature of management and agroclimatic variables on sweetpotato storage root yield variability. These phenology-driven models specified that storage root initiation did not start until 4 weeks after planting. However, we have documented that

storage root initiation, defined as the appearance of secondary cambium (Togari, 1950; Wilson and Lowe, 1973), can be detected as early as 13 d after transplanting (DAT) in 'Beauregard' grown under field conditions (Villordon et al., 2009a). The temporal discrepancy associated with the onset of storage root initiation may be attributable in part to differences in cultivar, management practices, prevailing environment, and the definition of "storage root initiation," i.e., visible (enlarged storage roots) versus anatomical (appearance of anomalous cambia in adventitious roots without visible enlargement).

The need for accurate prediction, inference, risk analysis, and decision-making is magnified by the underground nature and well-documented yield variability of sweetpotatoes. For example, Louisiana statewide total yields averaged 18.9 ton/ha in the last 20 years (NASS, 2009). Assuming 70% were graded as U.S. #1 storage roots, the calculated average U.S. #1 yield (US1YIELD) would have been 13.2 ton/ha or ≈1.2 U.S. #1 storage roots per hill (0.3-m in-row spacing, 1-m centers, average weight of roots = 0.34 kg, no. of plants per ha = 31,250). Yet, the US1YIELD potential of 'Beauregard' is clearly higher, i.e., 19.2 ton/ha (≈1.9 U.S. #1 per hill) (Rolston et al., 1987). The availability of well-calibrated agroclimatic-driven crop models will enable researchers to account for yield variability resulting from agroclimatic background and facilitate the investigation of effects attributed to management variables. This underscores the need to explore other modeling paradigms as well as to perform calibration of models for a cultivar grown under a specific production environment. Ji et al. (2007) suggested that a well-calibrated empirical model offers a more reliable method of investigating crop response than an inadequately calibrated process-based model.

An alternative modeling approach uses Bayesian belief networks (BBNs). BBNs graphically and probabilistically represent correlative and causal relationships among variables (Cain, 2001; Neopolitan, 2003). BBNs offer many advantages in modeling a domain (sweetpotato production system) that is characterized by an incomplete understanding of the interaction of agroclimatic, management, and biological variables. The graphical nature of BBNs facilitates ease in interpretation of the causal relationships among variables. In addition, BBNs allow combining of domain-specific (expert knowledge) and empirical data obtained from planned or ongoing experiments. An added benefit is that BBNs can learn from small and incomplete data sets (Uusitalo, 2007). BBNs have been used extensively in ecology and wildlife management to describe the influence of environmental variables on ecological-response variables (Marcot et al., 2006). In agriculture, BBNs have been used to model the effect of climate change in potato production (*Solanum tuberosum*) (Gu et al., 1994), predict yield response of winter wheat (*Triticum aestivum*) to fungicide application programs (Tari, 1996), and the development of a decision support system for growing malting barley (*Hordeum vulgare* L.) without the use of pesticides (Kristensen and Rasmussen, 2002).

Togari (1950) provided evidence that management and environmental variables within the first 20 DAT influenced adventitious root cambium activity in sweetpotato, which in turn determined the degree of lignification. It was determined early on that lignification rendered an adventitious root incapable of becoming a storage root and that storage root number was determined within this time period (Togari, 1950; Wilson

and Lowe, 1973). This timeframe for storage root initiation was locally validated (Chase, LA) for the sweetpotato cultivar Beauregard and that adventitious roots initiated within 7 DAT comprised 86% of final storage root yield under certain conditions (Villordon et al., 2009a, 2009b). Togari (1950) enumerated the following variables that increased cambium activity: potash, soil moisture, and optimum temperature. On the other hand, the following variables reduced cambium activity that eventually led to lignification: small seed roots, poor quality of transplants, suboptimum temperature, dry and compact soil, shading, and excessive nitrogen. The objective of this study was to develop a prototype or beta-level Bayesian network model to represent the relationship between agroclimatic variables measured within 20 DAT and U.S. #1 sweetpotato yield from experimental plots under uniform management conditions. We assumed the absence of influence from weeds, disease, insect pests, and chemical injury. The long-term goal is to develop a more comprehensive systemwide model that incorporates other management variables and identifies specific scenarios that result in economic loss for producers. We undertook this work using the rapid prototyping strategy (Connell and Shafer, 1989). Rapid prototyping is a development paradigm in which a scaled-down system or portion of a system is constructed in a relatively short period of time, tested, and improved through numerous iterations (Turban, 1992).

Materials and Methods

Study site. All field experiments were carried out from 2007 to 2009 in well-drained research fields in Chase, LA (lat. 32°6′ N, long. 91°42′ W). The soil taxonomic class

was fine-silty, mixed, active, thermic Typic Glossaqualfs. The soil bulk density was 1.3 g·cm^3 (15.2-cm depth, mean of three measurements). Virus-tested 'Beauregard' G1 seed roots (5.1 to 8.9 cm diameter and 7.6 to 22.9 cm in length) served as the transplant source and were bedded on the following dates in each of 3 years: 23 Mar. 2007, 31 Mar. 2008, and 3 Apr. 2009. In each year, seed roots were preheated at 80 °F for 7 d before bedding. Seed beds were fertilized with 50 kg·ha^{-1} nitrogen, 118 kg·ha^{-1} P$_2$O$_5$, and 135 kg·ha^{-1} K$_2$O. Research plots were prepared by disk cultivating fields followed by broadcast application of chlorpyrifos insecticide (2.2 kg ai/ha). A second disking operation was performed and then rows were formed on 1-m centers. In 2007, plots were 6 m long, 3 m in 2008, and 3 and 1.5 m (microplot) in 2009. In all years, record rows were separated by two guard rows. Plots were fertilized with 118 kg·ha^{-1} P$_2$O$_5$ and 135 kg·ha^{-1} K$_2$O. Nitrogen (50 kg·ha^{-1}) was side-dressed at 30 to 35 DAT or split-applied (half was applied at preplant and the remainder was side dressed at 30 to 35 DAT). Herbicide application, consisting of a tank mix of clomazone (840 g·ha^{-1}) and flumioxazin (36 g·ha^{-1}), was performed immediately before transplanting. In all years, cutting of transplants and setting was performed by one person to reduce variability associated with transplant operations, especially in transplant selection and depth of setting. All transplant operations were performed within 2 d of cutting planting materials. Transplants (eight to 12 nodes with intact leaves; at least 0.3 cm thick at the base) were cut from seed beds and planted on the same day or held upright overnight. The transplanting dates ranged from 15 May to 28 June across years. A white-skinned, white-fleshed cultivar (O'Henry) was used as a guard plant on each end of the plot to reduce the occurrence of oversized roots associated with the lack of intrahill competition in these locations. 'O'Henry' is a mutant selection of 'Beauregard' and did not represent any competitive difference. Uniform transplants were manually set (three to five nodes under the surface; in-row spacing = 30 cm) and watered in with ≈177 mL of water. Within 3 to 5 DAT, supplemental overhead irrigation was supplied with a traveling irrigation sprinkler if a rainfall event did not occur. This was routinely performed to prevent transplant desiccation and to help ensure uniformity of transplant establishment. Plant stand was determined 15 DAT (100% in all years). In all years, soil moisture was measured with a HydroSense Soil Water Content Management System (CS-620, 20-cm probe; Campbell Scientific, Inc. Logan, UT) that was calibrated gravimetrically. For the soil type used in the study, volumetric water content (VWC) = 16% represented ≈50% of field capacity and VWC in the range of 10% to 20% has been previously locally validated as optimal for sweetpotatoes grown in the area (Constantin et al., 1974). For transplant establishment, overhead irrigation was applied until ≈50% of field capacity. Thereafter, supplemental irrigation was supplied when soil mois-

ture approached 8% to 10% VWC at the 15-cm depth. After 70 DAT, supplemental irrigation was only applied when soil VWC approached 10%. Preharvest irrigation was only performed in extremely dry conditions to facilitate harvest operations. In 2009, soil moisture monitoring was augmented by the installation of soil moisture sensors (ECH2O EC-5; Decagon Devices Inc., Pullman, WA) linked to automated data loggers (EM50; Decagon Devices Inc.). The data loggers were connected wirelessly to a data collection device (DataStation; Decagon Devices Inc.). Soil moisture sensors (5 cm in length) were installed vertically at two depths (5 and 15 cm) in two plots in each of three planting dates. The purpose of these automated sensors was to better document soil moisture variation in the soil profile (Fig. 1). The ECH2O sensor readings were calibrated with the CS-620 readings. For the purpose of this work, we used the VWC readings obtained from the CS-620 20-cm probe.

We have previously described a simple system of describing morphologically distinct storage root initiation phenological stages in 'Beauregard': SR1 (transplant establishment, i.e., terminal leaves start to open, appearance of adventitious roots, protoxylem development, primary cambium development) (Fig. 2A), SR2 (seven to 14 new leaves; one to four branches; appearance of anomalous or secondary cambium; storage root initiation) (Fig. 2B–C), and SR3 (21 to 42 new leaves; three to four branches; appearance of visible initiated roots; adventitious root with visible localized swelling, minimum of 0.5 cm in diameter or greater) (Fig. 2D) (Villordon et al., 2009b).We have adopted this approach to describe the presumptive phenological stages in this work. In 2008 and 2009, five to 10 random plants in each planting date were sampled to verify onset of SR1 (3 to 7 DAT) and SR2 (13 to 18 DAT). In SR2 sampling, all adventitious roots (greater than 5 cm long) from each plant were sectioned, stained with toluidine blue (Eguchi and Yoshida, 2008), and examined under a microscope. A similar sampling approach was also used to verify SR3 (28 to 35 DAT). In 2007 and 2008, a one-row mechanical digger was used for harvesting. In 2009, plots were manually harvested as a result of successive rainfall events in August, September, and October. The harvest dates ranged from 13 Aug. to 7 Nov. across years. At harvest, storage roots were graded according to U.S. Department of Agriculture (USDA) standards (USDA, 2005): U.S. #1 (5.1 to 8.9 cm diameter and 7.6 to 22.9 cm in length), canner (2.5 to 5.1 cm in diameter and 5.1 to 17.8 cm in length), and jumbo (larger than both groups). Storage roots were counted and weighed. To reduce variability associated with grading of yield classes, only one person performed grading and counting in all years, and digital photographs were taken of all modeling plots for later reference. Leaf tissue samples (fifth fully expanded leaf from the terminal of the longest branch) were collected 35 to 50 DAT from one representative planting date in each of 2007 and 2008.
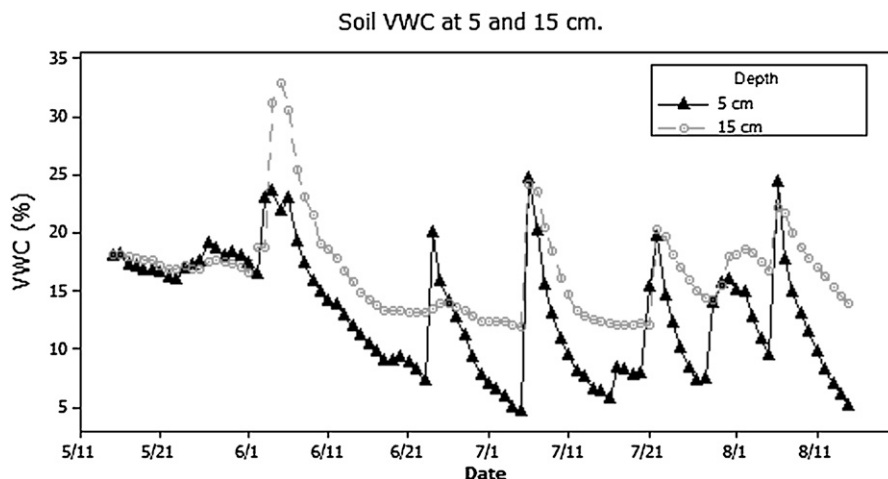
Fig. 1. Plots of soil moisture, expressed as volumetric water content (VWC), from 'Beauregard' sweetpotato field plots in Chase, LA, in 2009. ECH2O EC-5 sensors (Decagon Devices Inc., Pullman, WA) were installed vertically at 5 and 15 cm. The sensors were connected to a data collection device (EM50 data logger; Decagon Devices Inc.) wirelessly connected to a base radio station (DataStation; Decagon Devices Inc.). Each rise in VWC represented a rainfall or irrigation event.
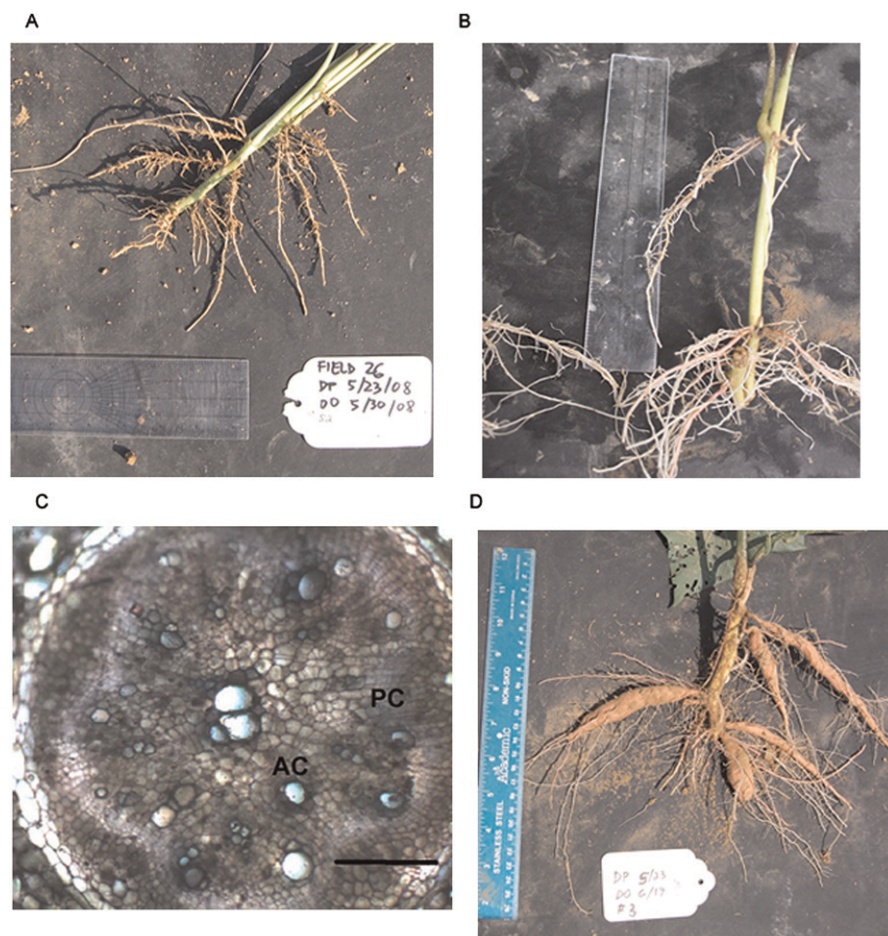

Fig. 2. Representative adventitious root samples showing presumptive phenological stages SR1 (**A**), SR2 (**B–C**), and SR3 (**D**) of 'Beauregard' sweetpotato grown in Chase, LA. 1C depicts the appearance of anomalous cambium (AC) at SR2. PC = primary cambium; scale bar = 200 uM. SR1 sample was obtained from plots transplanted 23 May 2008 and sampled 30 May 2008. SR2 sample was obtained from plots transplanted 15 May 2009 and sampled 29 May 2009. SR3 sample was obtained from plots transplanted 23 May 2008 and sampled on 19 June 2008. Adventitious roots from 1B were used as source of cross-sections for 1C. Agroclimatic and management variables are described in "Materials and Methods."

Soil samples were collected in November in each of 3 years. Leaf and soil samples were analyzed at the LSU AgCenter Soil Test and Plant Analysis Laboratory, Baton Rouge, LA. Soil analysis included organic matter, pH, phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), sodium (Na), sulfur (S), copper (Cu), zinc (Zn), and boron (B). Leaf tissue analysis included P, K, Ca, Mg, S, B, Zn, Cu, Fe, Manganese, aluminum, Na, and total nitrogen. Description of analytical procedures is available from http://www.lsuagcenter.com/stpal/ (1 Jan. 2010). Results of the soil and plant tissue analyses did not show any significant deviation from currently acceptable levels of soil and plant tissue nutrient ranges (Bouwkamp, 1985; Mills and Jones, 1996) (data not shown). Furthermore, nematode assays showed that nematode damage potential was classified as "low" in all years by the LSU AgCenter Nematode Advisory Service (data not shown).

Agroclimatic data were obtained from on-site National Oceanic and Atmospheric Administration (soil and air temperatures, rainfall) and Louisiana Agriclimatic Information System (relative humidity, solar radiation) weather stations, respectively. Descriptive statistics of agroclimatic information from May to September in each year are presented in Figure 3.

Data preprocessing. There were 21 records (21 planting and harvest date combinations) in the modeling data set (MDS). MDS variables included U.S. #1 count (US1COUNT) and weight, growing degree-days (GDD) for each planting date–harvest date combination, and agroclimatic data. We have previously empirically derived a method for calculating GDD for Louisiana-grown sweetpotatoes (Villordon et al., 2009c). Briefly, GDD was calculated using this method: maximum daily temperature (Tmax) – base temperature (B, 15.5 °C), where if Tmax > ceiling temperature (C, 32.2 °C), then Tmax = C, and where GDD = 0 if minimum daily temperature (Tmin) < B. GDD was used to represent accumulated air heat units in each phenological phase and to adjust for differences in growing periods (planting and harvest date combinations). Soil temperature was represented as soil heat units (SHU). SHU was calculated using this method: $\{[(Tmax + Tmin)/2] - B\}$ where SHU = 0 if Tmin < B (B = 18.3 °C). The method for calculating SHU as well as C and B was empirically derived from the MDS using the minimum cv approach (Dufault, 1997; Jenni et al., 1996). The data were entered into an Excel spreadsheet (Version 7; Microsoft, Redmond, WA) and natural log transformed. Transformation was performed to make the estimation process more robust and reduce unstable results (Kuhnert and Hayes, 2009). Correlation analysis was performed among the agroclimatic variables to determine if linkages were going to be established during model development (Marcot et al., 2006).

Development paradigm. We adopted the general guidelines that Marcot et al. (2006)
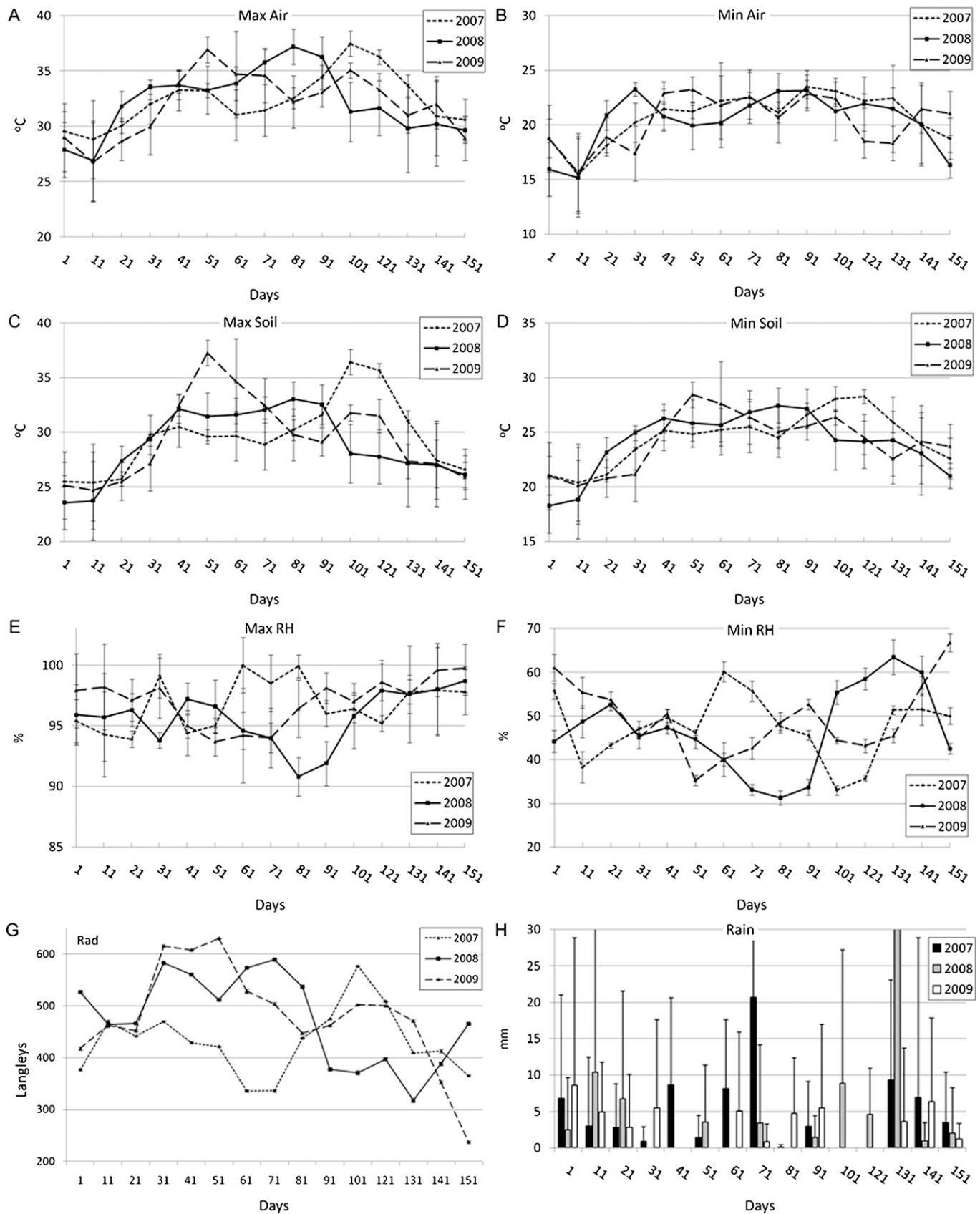
Fig. 3. Plots of mean values of air temperature (**A–B**), soil temperature (**C–D**), relative humidity (**E–F**), solar radiation (**G**), and rainfall (**H**) during most of the growing season for 'Beauregard' sweetpotato grown in Chase, LA (2007 to 2009). Values represent means calculated at 10-d intervals starting from 1 May to 30 Sept. in each year. Agroclimatic data were obtained from on-site National Oceanic and Atmospheric Administration (NOAA) (soil and air temperatures, rainfall) and Louisiana Agriclimatic Information System (relative humidity, solar radiation) weather stations. Soil temperature depth = 15 cm. Bars = SD (n = 10). Max = maximum; min = minimum; Air = air temperature (°C); soil = soil temperature (°C); RH = relative humidity (%); Rad = solar radiation (Langleys; 1 Langley/d = 0.48 W·m⁻²); Rain = rainfall (mm).

proposed for developing, testing, and revising BBNs, especially in identifying alpha-, beta-, and gamma-level models. The first step comprised of creating influence diagrams (AM1 to AM12) of hypothesized "causal web" of agroclimatic variables affecting the outcome of interest (US1YIELD) (Fig. 4). The influence diagrams were constructed using Netica software (Version 4.09; Norsys Software Corp., Vancouver, Canada). The underlying biological principle in the development of the model structure was that storage root number (US1COUNT) was determined in the first 20 DAT and that final weight at harvest (US1YIELD) was a function of time [days to harvest (H) expressed as GDDH] assuming other variables were uniform. The first 20 DAT was also classified into two distinct phenological stages: SR1 (1 to 10 DAT) and SR2 (10 to 20 DAT) (Villordon et al., 2009a). Where present, SR1 and SR2 were treated as "latent nodes" or "hidden variables." The network structures or topologies of candidate alpha-level models AM1, AM3, and AM4 represented the hypothetical relationship of agroclimatic variables and US1YIELD assuming the absence of SR1 and SR2 effects. The network

structures of AM1 and AM2 resembled naïve Bayes topology and were used to represent models that presumed the lack of dependencies among the causal variables. AM1, AM3, and AM4 used agroclimatic data for the first 20 DAT. In contrast, the network topologies of AM2 and AM5 to AM9 assumed the existence of phenological stages SR1 and SR2 in helping to represent the relationship between agroclimatic variables and US1YIELD. The hypothetical network structures of several candidate models (AM5 to 10) were based in part on a published phenology-driven BBN model for potato growth and development (Gu et al., 1994). Where present, arcs or links among agroclimatic variables were determined through correlation analysis (Marcot et al., 2006). AM11 was a structure derived from the "learn new network" function of GENiE (Bersion 2.0; Decision Systems Laboratory, University of Pittsburgh, Pittsburgh, PA). In GENiE's network structure learning algorithm, agroclimatic predictor variables were assigned to temporal tiers and US1YIELD was specified as dependent on GDDH and US1COUNT. The learning method was greedy thick thinning. AM12 was a structure derived from the "learning wizard" function

of Hugin Researcher (Version 7.2; Hugin Expert A/S, Aalborg, Denmark). In this mode, Hugin Researcher attempted to learn the model structure without prior assumptions of any relationships among all variables in the data set. The only supplied inputs (final step of "learning wizard") were to ensure that applicable variables influenced US1COUNT or US1YIELD. As a result of current limitations of GENiE and Hugin Researcher regarding the use of latent variables with missing values, SR1 and SR2 were not included in the identification of the network structure. In the second step, alpha-level BBNs were subsequently developed from the hypothetical influence diagrams through parameter learning from empirical data, i.e., MDS. A candidate beta-level BBN was subsequently selected from these alpha BBNs based on estimates of accuracy and model performance obtained through cross-validation, validation on independent data set, receiver operating characteristic (ROC) analysis, and performance relative to a baseline model (logistic regression).

Parameter learning. The procedures for parameter learning [determination of conditional probability table (CPT) at each node] from the experimental data set in a spreadsheet
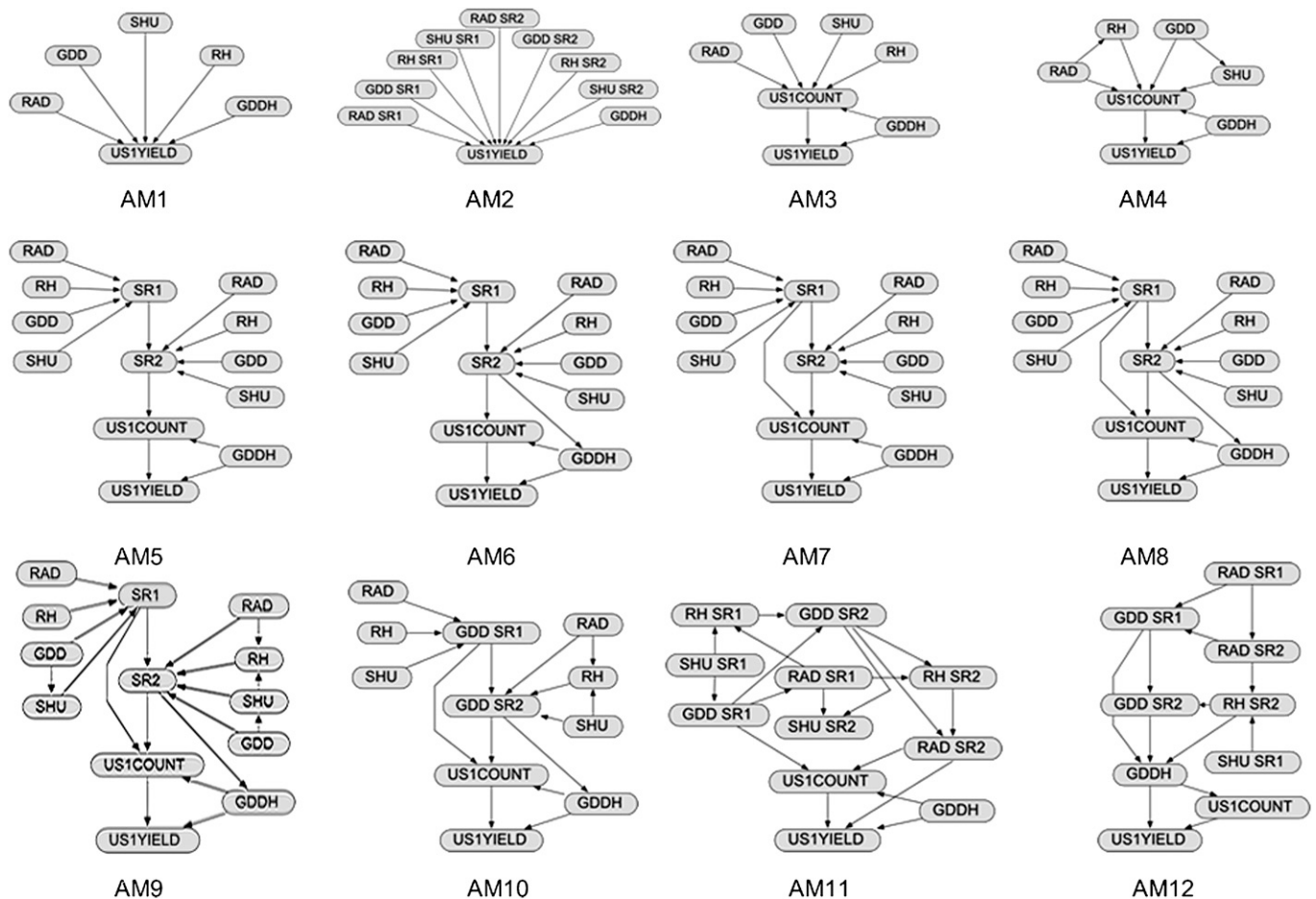
Fig. 4. Alpha-level Bayesian belief network structures evaluated for prototype beta-level model development of the relationship between U.S. #1 yield and agroclimatic variables known to influence storage root initiation in sweetpotato 'Beauregard' grown in Louisiana. RAD = solar radiation (Langleys; 1 Langley/d = 0.48 W·m^{-2}); RH = relative humidity (%); GDDH = growing degree-days to harvest; SHU = soil heat units; US1COUNT = U.S. #1 count; US1YIELD = U.S. #1 yield in tons/ha. Where present, SR1 and SR2 nodes represented presumptive phenological stages corresponding to protoxylem development (1 to 10 DAT) and anomalous cambium development or storage root initiation (10 to 20 DAT), respectively.

file format were based on the procedures described in Netica's built-in user manual. The learning algorithm was gradient descent (Binder et al., 1997) as implemented in Netica. This algorithm permits learning when some variables are hidden (Binder et al., 1997), i.e., variables that are not currently measurable (SR1 and SR2). Discretization was performed using Netica's built-in discretization algorithm. Except for US1YIELD and US1COUNT, all nodes were discretized into three states. US1YIELD and US1COUNT were discretized into two states. Preliminary modeling experiments showed that three-state discretization of US1YIELD and US1COUNT resulted in poor performing models (data not shown).

Model assessment and selection. The comparative performance of the BBN models was estimated using the following approaches: cross-validation (leave-one-out), validation with independent (unseen) test data set (VID), and area under the receiving operator characteristic curve (AUC) analysis of the VID results. In leave-one-out cross-validation, given $n = 21$ MDS data records, n-1 was used to learn the model and the remaining one record was used to test model performance. The procedure was repeated n times, each time with different learning–testing record combinations. For each model, the error rate was reported as the percentage of misclassified cases across n runs. VID was performed with data that were previously used to develop a GDD model (Villordon et al., 2009c). A subset of this data set (n = 98), comprised of yield and agroclimatic data from 2002 to 2006, was used for validation. This data set comprised yield data from replicated tests in research plots as well as grower fields. As a result of limitations of earlier experimental methodology, this test data set did not contain soil moisture data. To simulate replications (pseudoreplications), Insightful Data Miner (Version 8; Insightful Corp., Seattle, WA) was used to randomly generate 10 unique data partitions (n = 50% of the test data set). Each candidate alpha BBN was subsequently validated on each pseudoreplicate ($r = 10$). Measurements of accuracy included error rate and AUC analysis (mean of 10 pseudoreplicates). These values were calculated based on the "test with cases" option in Netica to classify the outcome of unseen data. Error rate was automatically calculated by Netica when performing the "test with cases" algorithm. This value (expressed in percentage) was calculated by dividing the number of misclassified cases by the total number of classifications made. For AUC analysis, the "test with cases" option in Netica generated sensitivity and specificity estimates as described in Netica's built-in user manual. The sensitivity and 1-specificity estimates were subsequently entered into a ROC analysis spreadsheet (Watkins, 2000) for calculation of AUC values and for plotting the ROC curve. The predictive accuracy of the candidate alpha BBNs were compared with logistic regression (baseline model). U.S. #1 yield was reclassified into "LOW" or "HIGH" based on a threshold of

24 tons/ha US1YIELD (LOW = US1YIELD < 24 tons/ha; HIGH = US1YIELD > 24 tons/ha). This threshold was based on the discretization threshold for US1YIELD in the BBN model. Logistic regression analysis using forward selection was performed in SAS Analyst (Version 9.2; SAS Institute, Cary, NC). GDD-related variables were manually included where feasible, i.e., convergence criterion was satisfied. ROC curves from the logistic regression analyses were calculated using the ROC Curve procedure in SPSS (Version 15.0; SPSS Inc., Chicago, IL). The performance of the candidate models were then ranked based on cross-validation performance, VID error rates, and AUC analysis.

Results

There was agroclimatic variability within and among periods corresponding to storage root number determination (SR1 and SR2) and initial enlargement (SR3) within and among years (Fig. 3). For example, in 2007, air temperatures peaked around mid-August (101 d; Fig. 3A–B). However, the peak air temperatures in each of 2008 and 2009 were achieved in late July (81 d) and late June (51 d), respectively (Fig. 3A–B). In each year, soil temperature varied considerably among planting dates (Fig. 3C–D). For example, in 2009, the mean minimum soil temperature during the last 14 d of May (16 to 31 d) and $\approx$10 d into June (32 to 42 d) was 21 °C; in the next 14 d (43 to 57 d), this value was 26 °C (Fig. 3D). There was a wide divergence of measured relative humidity (RH) across years, especially around the start of the transplanting period (mid-May; 15 d) and in

late June to early July (46 to 76 d; Fig. 3E–F). Minimum RH was below 40% for certain periods in 2008 (mid-July; 71 to 91 d) and 2009 (late June; 51 to 61 d) (Fig. 3F). Measured solar radiation showed distinct variability within and across years (Fig. 3G). Mean daily net solar radiation during most of June (32 to 61 d) was comparatively lower in 2007 (below 500 Langleys/d or 242 W·m^{-2}) than in 2008 and 2009 (Fig. 3G). Rainfall events varied among years, although rainfall distribution was relatively more uniform during the transplanting period in 2007 (mid-May to late June; 15 to 61 d) (Fig. 3H). There were no rainfall events recorded for certain periods in 2008 (between 29 May and 19 June; 29 to 50 d) and 2009 (between 5 June and 30 June; 36 to 61 d) (Fig. 3H). Model development was based on the assumption that other variables such as soil nutrients, crop nutrition, soil moisture, disease, etc., were not limiting. In 2009, further description of soil moisture states at two depths (5 and 15 cm) showed intraseason soil moisture variability (Fig. 1). In general, the soil moisture variability within the soil profile, as influenced by rainfall and irrigation events, was more pronounced when air and soil temperatures began to increase during the transplanting period, e.g., after the first week of June in 2009.

Candidate model AM9 consistently ranked high across the various measures of model accuracy (Table 1). AM9 was the only model structure that also consistently showed comparable performance relative to the logistic regression model across all measures of accuracy. The main difference between AM9 versus the other candidate models was the

Table 1. Comparative predictive performance of alpha-level Bayesian belief network and logistic regression models using leave-one-out cross-validation error rates, validation on independent test data set error rates, and area under the receiver operating characteristic curve (AUC) for representing the relationship between agroclimatic predictor variables and sweetpotato U.S.#1 yield in Louisiana.

Model[z]	Cross-validation error rate[y] (%)	Validation[x]			
		Error rate (%)	SE	AUC	SE
AM1	19	42 d	1.2	0.30 e	0.02
AM2	29	46 d	2.1	0.62 a	0.02
AM3	14	61 f	2.1	0.25 e	0.02
AM4	14	45 d	2.0	0.29 e	0.03
AM5	14	55 e	1.6	0.39 d	0.02
AM6	14	54 e	1.3	0.42 cd	0.02
AM7	14	43 d	1.6	0.45 bc	0.02
AM8	14	23 b	1.4	0.41 cd	0.02
AM9	14	20 ab	1.2	0.59 a	0.02
AM10	14	53 e	2.6	0.37 d	0.03
AM11	14	30 c	1.4	0.58 a	0.03
AM12	14	17 a	2.3	0.50 b	0.00
LOGISTIC	14	18 a	1.5	0.62 a	0.02

[z]AM1 to AM12 represented candidate Bayesian belief network topologies defined in the "Materials and Methods" section. LOGISTIC = logistic regression (forward selection) was performed in SAS Analyst (Version 9; SAS Institute, Cary, NC). LOGISTIC was considered as the baseline model. Validation error rates represent mean (10 pseudoreplicates) error rates ± SE of the mean.
[y]Cross-validation method was leave-one-out where n = 21. For each model, error rate was calculated as the percentage of misclassified cases across n runs.
[x]The independent data set (n = 98) was partitioned into 10 unique data partitions (n = 50%) and each model was run on all partitions (10 pseudoreplicates). Error rates were calculated with the Netica "test with cases" option. Calculation of AUC is defined in the "Materials and Methods" section. AUC values represent mean (10 pseudoreplicates) error rates. SE = standard error of the mean. Means and mean rank followed by the same letter within columns are not significantly different at $P < 0.05$. Underscored mean error rates and AUC represented model performance estimates that were not significantly different relative to LOGISTIC.

presence of additional arcs or links that represented correlations among causal variables and the causal influence of SR1 on final yield. AM9 was chosen as the final prototype model (beta-level BBN). The compiled AM9 model (here after referred to as BxNET version 1.00b) is shown in Figure 5 and the beliefs are shown for each node in the form of belief bars. 'Beauregard' has frequently been coded as "BX" in field plots, genotype screening trials as well as in some published reports, for example Mcharo and LaBonte (2007). The version number designation follows the recommendation by Marcot et al. (2006) in which 1.00b signifies beta-level status.

BxNET was used to simulate the outcomes of multiple scenarios between the predictor variables and US1YIELD (Fig. 6). Scenario 6A represented relatively low air and soil temperatures, similar to the conditions that were observed during the second half of May (15 to 31 d) in each of 2007, 2008, and 2009 (Fig. 3A–D). Under these conditions, mean air and soil temperature ranges were 17 to 28 °C and 20 to 25 °C, respectively. Scenario 6B represented relatively high soil and air temperatures, similar to a 3-week period in June 2009 (Fig. 3A–D; 41 to 61 d). During this period, the mean ranges of air and soil temperature were 22 to 36 °C and 27 to 36 °C, respectively. Scenario 6C represented "intermediate" solar radiation and air and soil temperatures, similar to conditions observed during the last week of May and first two weeks of June 2007 (Fig. 3A–D; 25 to 45 d). The "intermediate" mean air and soil temperature ranges were 20 to 31 °C and 23 to 8 °C, respectively. The finding for mean RH was entered as 71% to 73% in all cases. Given GDD = 1319 to 1383, the predicted output varied: US1YIELD = HIGH (US1YIELD > 24 t·ha⁻¹) probabilities for scenarios 6A, 6B, and 6C were 42%, 2%, and 57%, respectively. When GDD was increased to 1383 to 1579 (longer growing period), the predicted output varied: US1YIELD = HIGH probabilities for scenarios 6A, 6B, and 6C were 28% (Fig. 6D), 0.04% (Fig. 6E), and 99% (Fig. 6F), respectively. In general, GDD estimates varied with calendar days, but we were able to experimentally attain GDD = 1383 at 105 DAT. Thus, 1383 to 1579 GDD represented 105 to 119 DAT. Past calendar day-based studies indicated that 115 DAT was the preferred number of days to harvest for 'Beauregard' (Villordon et al., 2003).

Discussion

Sweetpotato crop yield components include plant density (plant spacing and stand), storage root count per hill, and weight. All of these variables were accounted for in the development of sweetpotato 'Beauregard' beta-level BBN named BxNET (Version 1.00b). This phenology-driven prototype model represented the relationship between US1YIELD and agroclimatic variables (air and soil temperature, RH, and solar radiation) that have been documented to directly
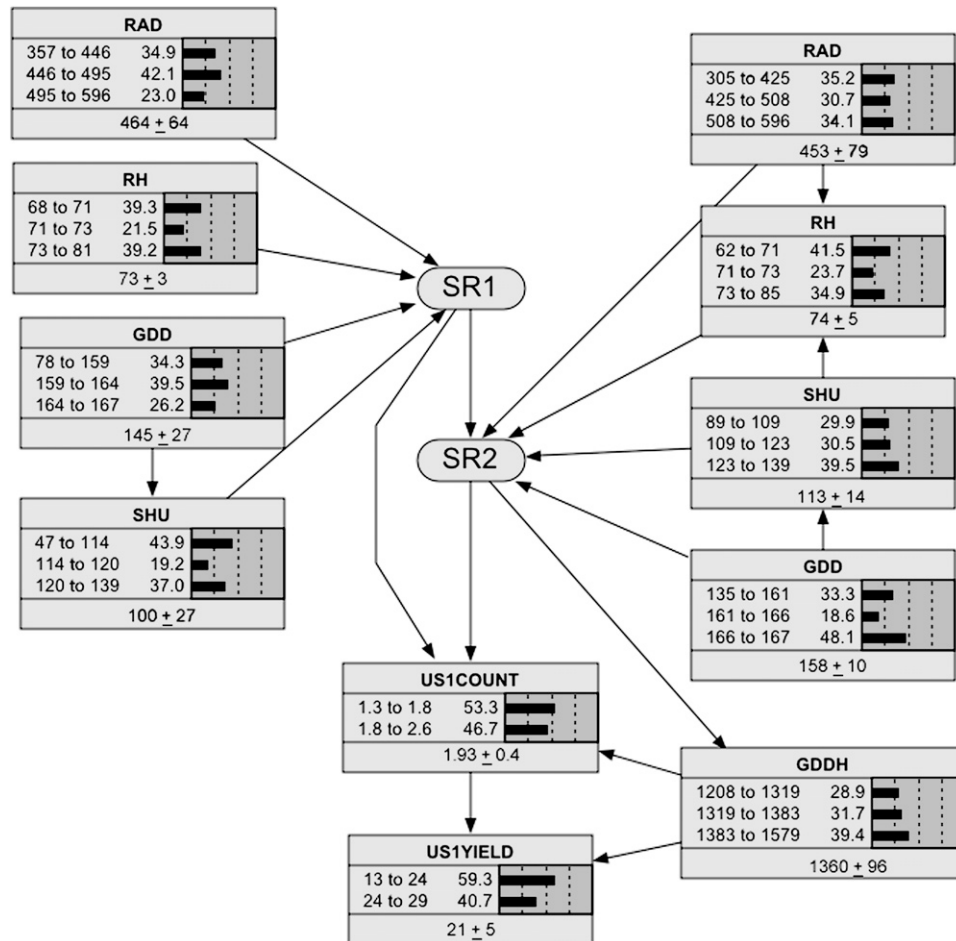
Fig. 5. A prototype Bayesian belief network model, named BxNET (Version 1.00b), representing casual effects of some agroclimatic variables on U.S. #1 yield in 'Beauregard' sweetpotato grown in Chase, LA. BxNET was trained using the gradient descent method as implemented in Netica (Version 4.09; Norsys Software Corp., Vancouver, Canada). Horizontal bars (belief bars) and values within nodes are probabilities of states of each variable. Natural log-transformed data were used to calculate the conditional probability values. Corresponding non-transformed ranges are shown for all variable states or ranges. Mean transformed value ± SD are shown below the belief bars. Intermediate nodes SR1 and SR2 represent presumptive phenological stages corresponding to protoxylem development (1 to 10 DAT) and anomalous cambium development or storage root initiation (10 to 20 DAT), respectively. Input nodes: RAD = solar radiation (Langleys; 1 Langley/d = 0.48 W·m⁻²); RH = relative humidity (%); GDDH = growing degree-days to harvest; SHU = soil heat units; US1COUNT = U.S. #1 count. Output node: US1YIELD = U.S. #1 yield in tons/ha. Details of the experimental data and other procedures are defined in "Materials and Methods."

A

RADSR1	
357 to 446	100
446 to 495	0
495 to 596	0
402 ± 26	

GDDSR1	
78 to 159	100
159 to 164	0
164 to 167	0
119 ± 23	

SHUSR1	
47 to 114	100
114 to 120	0
120 to 139	0
80 ± 20	

GDDH	
1208 to 1319	0
1319 to 1383	100
1383 to 1579	0
1351 ± 18	

US1COUNT	
1.3 to 1.8	36.6
1.8 to 2.6	63.4
2 ± 0.4	

RADSR2	
305 to 425	100
425 to 508	0
508 to 596	0
364 ± 34	

GDDSR2	
135 to 161	100
161 to 166	0
166 to 167	0
148 ± 7	

SHUSR2	
89 to 109	100
109 to 123	0
123 to 139	0
99 ± 6	

US1YIELD	
13 to 24	57.8
24 to 29	42.2
22 ± 5	

B

RADSR1	
357 to 446	0
446 to 495	0
495 to 596	100
545 ± 29	

GDDSR1	
78 to 159	0
159 to 164	0
164 to 167	100
166 ± 1	

SHUSR1	
47 to 114	0
114 to 120	0
120 to 139	100
130 ± 6	

GDDH	
1208 to 1319	0
1319 to 1383	100
1383 to 1579	0
1351 ± 18	

US1COUNT	
1.3 to 1.8	95.8
1.8 to 2.6	4.25
1.6 ± 0.2	

RADSR2	
305 to 425	0
425 to 508	0
508 to 596	100
552 ± 25	

GDDSR2	
135 to 161	0
161 to 166	0
166 to 167	100
167 ± 1	

SHUSR2	
89 to 109	0
109 to 123	0
123 to 139	100
131 ± 18	

US1YIELD	
13 to 24	97.2
24 to 29	2.82
19 ± 3	

C

RADSR1	
357 to 446	0
446 to 495	100
495 to 596	0
471 ± 14	

GDDSR1	
78 to 159	0
159 to 164	100
164 to 167	0
162 ± 1	

SHUSR1	
47 to 114	0
114 to 120	100
120 to 139	0
117 ± 2	

GDDH	
1208 to 1319	0
1319 to 1383	100
1383 to 1579	0
1351 ± 18	

US1COUNT	
1.3 to 1.8	14.2
1.8 to 2.6	85.8
2.1 ± 0.3	

RADSR2	
305 to 425	0
425 to 508	100
508 to 596	0
465 ± 25	

GDDSR2	
135 to 161	0
161 to 166	100
166 to 167	0
164 ± 1	

SHUSR2	
89 to 109	0
109 to 123	100
123 to 139	0
116 ± 1	

US1YIELD	
13 to 24	42.9
24 to 29	57.1
24 ± 5	

D

GDDH	
1208 to 1319	0
1319 to 1383	0
1383 to 1579	100
1482 ± 57	

US1COUNT	
1.3 to 1.8	71.8
1.8 to 2.6	28.2
1.7 ± 0.3	

US1YIELD	
13 to 24	71.8
24 to 29	28.2
22 ± 5	

E

GDDH	
1208 to 1319	0
1319 to 1383	0
1383 to 1579	100
1482 ± 57	

US1COUNT	
1.3 to 1.8	100
1.8 to 2.6	.042
1.5 ± 0.1	

US1YIELD	
13 to 24	100
24 to 29	.042
19 ± 3	

F

GDDH	
1208 to 1319	0
1319 to 1383	0
1383 to 1579	100
1482 ± 57	

US1COUNT	
1.3 to 1.8	0.73
1.8 to 2.6	99.3
2.2 ± 0.2	

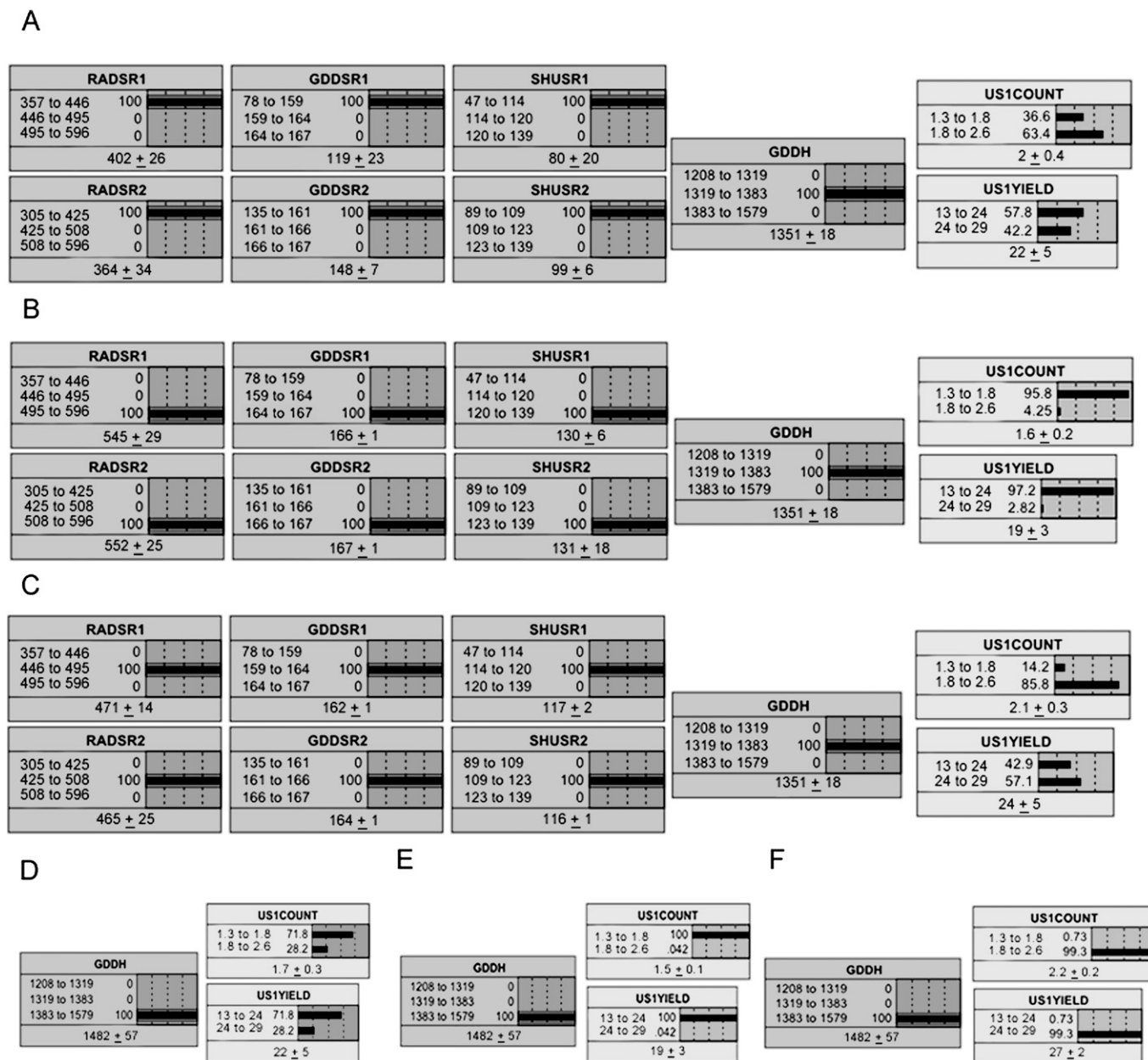| US1YIELD	
13 to 24	0.73
24 to 29	99.3
27 ± 2	

Fig. 6. Bayesian belief network scenarios representing the relationship between days to harvest [growing degree-days (GDD)] and U.S. #1 storage root yield as influenced by air and soil temperatures measured during SR1 and SR2 storage root initiation stages of 'Beauregard' sweetpotato grown in Louisiana. Scenario A represented relatively low air and soil temperatures during SR1 and SR2 (minimum air and soil heat unit accumulation). Scenario B represented relatively high air and soil temperatures during SR1 and SR2 (maximum air and soil heat unit accumulation). Scenario C represented intermediate low air and soil temperatures during SR1 and SR2 (intermediate air and heat unit accumulation). Scenarios D, E, and F represented extension of growing period for Scenarios A, B, and C, respectively. Days to harvest was expressed as growing degree-days (GDDH). The finding for RH (not visible) was entered at 71% to 73% and soil moisture ranged from 10% to 20% for all scenarios. Details of the experimental data and other procedures are defined in "Materials and Methods." Natural log-transformed data were used to calculate the conditional probability tables. Corresponding non-transformed ranges are shown for all variable states or ranges. Mean non-transformed value ± SD are shown below the belief bars. SR1 and SR2 represent presumptive phenological stages corresponding to protoxylem development (1 to 10 DAT) and anomalous cambium development or storage root initiation (10 to 20 DAT), respectively. RAD = solar radiation (Langleys; 1 Langley/d = 0.48 W·m^{-2}); RH = relative humidity (%); GDDH = growing degree-days to harvest; SHU = soil heat units; US1COUNT = U.S. #1 count; US1YIELD = U.S. #1 yield in tons/ha.

influence the determination of storage root number (Eguchi et al., 1998; Togari, 1950). The presumptive phenological stages were represented by the latent nodes SR1 and SR2. Within the context of BBN development, latent or intermediate nodes are typically used to summarize major themes in influence diagrams (Marcot et al., 2006). Within this modeling paradigm, we used SR1 and SR2 to represent hypothetical biological response to external stimuli. If detected, such nodes or variables that currently cannot be measured might lead to better understanding of the domain under consideration or even to scientific discovery (Zhang et al., 2004). As an example of its potential use, BxNET was used to simulate the modulating influence of prevailing SR1 and SR2 agroclimatic conditions on the relationship between days to harvest (expressed as GDD) and US1YIELD (Fig. 6). The simulation scenarios demonstrated that delaying harvest did not always lead to increased U.S. #1 yield; rather, the outcome was determined in part by the modulating effects of agroclimatic conditions extant at SR1 and SR2. The likelihood of obtaining US1YIELD = HIGH was greatest (probability = 99%) when air and soil heat unit accumulation were moderate during SR1 and SR2 (Figs. 6C and 6F). When the

prevailing air and soil temperatures were relatively low (minimum air and soil heat unit accumulation) at SR1 and SR2, the probability for US1YIELD = HIGH decreased to 28% (Fig. 6D). When air and soil temperatures were relatively high during SR1 and SR2 (maximum air and soil heat unit accumulation), the probability for obtaining US1YIELD = LOW increased to almost 100% (Fig. 6E). As an example, we observed that US1COUNT and US1YIELD were 1.9 and 20 tons/ha, respectively, under low prevailing air and soil temperatures at SR1 and SR2 (planting date = 15 May 2009; GDD SR1 = 78; GDD SR2 = 135; SHU SR1 = 46; SHU SR2 = 89; GDDH = 1335) (data not shown). Under relatively high prevailing air and soil temperatures at SR1 and SR2, we observed that US1COUNT and US1YIELD were 1.7 and 22.3 tons/ha, respectively (planting date = 15 June 2009; GDD SR1 = 167; GDD SR2 = 167; SHU SR1 = 139; SHU SR2 = 139; GDDH = 1346) (data not shown). In contrast, we observed that US1COUNT and US1YIELD were 2.5 and 29 tons/ha, respectively, when SR1 and SR2 occurred under "moderate" air and soil temperatures (planting date = 30 May 2007; GDD SR1 = 159; GDD SR2 = 164; SHU SR1 = 113; SHU SR2 = 119; GDDH = 1375) (data not shown). These results are consistent with past observations that sweetpotato storage root formation and growth are sensitive to relatively low and high air and soil temperatures (Bouwkamp, 1985; Ravi and Indira, 1999). Past research has shown that the yield of low temperature-treated (10 to 15 °C) plants was only 75% of controls (20 to 25 °C), although the former showed comparable vine growth when taken out of the low-temperature treatment (Bouwkamp, 1985). Soil temperatures between 20 and 30 °C favored storage root formation, whereas soil temperatures greater than 30 °C generally promoted shoot growth

at the expense of storage root growth (Ravi and Indira, 1999). To help explain the results of these modeling scenarios, representative storage root yields from various plots in 2009 are shown in Figure 7A–B. From a biological perspective, the majority (86%) of adventitious roots that were initiated from the underground nodes at 3 to 7 DAT (SR1, Fig. 2A) possessed the potential to become storage roots (Villordon et al., 2009a). Storage root initiation, defined as the appearance of anomalous or secondary cambium, occurred 13 to 18 DAT under field conditions (SR2, Fig. 2B–C). When SR1 and SR2 occurred within this timeframe (3 to 20 DAT), initiated storage roots typically attained a diameter of 0.5 cm or more at 30 DAT (SR3, Fig. 2D). Under "moderate" conditions, soil moisture was generally uniform as measured at the 5- and 15-cm depths (Fig. 1) and storage root initiation was observed across all underground nodes (Figs. 2D and 7A). In contrast, storage roots were generally initiated at the lower nodes (Fig. 7B) when relatively high air and soil temperatures prevailed during SR1 and SR2. Such conditions interacted with soil moisture at the 5-cm depth that approached the presumptive soil moisture threshold (8% to 10% VWC; Fig. 1) that did not favor consistent storage root initiation for the soil type used in the study.

Our results are consistent with past observations regarding the influence of solar radiation, RH, and air and soil temperature on sweetpotato growth, development, and storage root yield. These variables were also used as the main "driving" variables in process-based models developed for sweetpotato (Mithra and Somasundaram, 2008; Somasundaram and Mithra, 2008). Togari (1950) provided the necessary anatomical evidence that the growing environment 20 DAT directly influenced cambium activity, which in turn influenced the subsequent development of storage roots. Kays (1985) reviewed the influence of

environmental factors on sweetpotato yield and indicated that photosynthesis of individual leaves saturated at approximately one-third of full sunlight. Quality of radiant energy has not been well studied in sweetpotatoes attributable in part to the limited degree of influence over it (Kays, 1985). Eguchi et al. (1998) experimentally demonstrated that storage roots attained maximum growth at 70% RH versus 50% and 90% RH when air temperature was maintained at 27 °C. Eguchi et al. (1994) also experimentally demonstrated a curvilinear response of storage root weight to sink (soil) temperature with 23 to 26 °C being the optimal range when air temperature (27 °C) and RH (70%) were constant.

The previous example of the potential use of BxNET represents the use of a BBN in a limited predictive mode. A BBN can also be used in diagnostic or inference mode. For example, if in future validation work BxNET predicts US1YIELD = HIGH when in fact it is LOW (i.e., a "false-positive"), then the possible causes of the unexpected suboptimal yield can be systematically investigated. If agroclimatic variables, transplant quality, soil moisture, soil fertility, crop nutrition, poor plant stand, harvest date, and weed and disease effects can be ruled out, then other possible causes can be considered. Examination of the yield outcome will also help in providing clues. If the poor yield is the result of the reduced proportion of storage roots relative to lignified or "stringy" roots across all underground nodes, then an accidental or an unplanned intervention event at SR1 or SR2 cannot be ruled out. Recent research data have implicated that sucrose and cytokinin (Eguchi and Yoshida, 2008; Tanaka et al., 2008) are necessary for storage root initiation. The major active endogenous cytokinin in storage roots has been identified as transzeatin (Tanaka et al., 2008). Eguchi and Yoshida (2008) experimentally demonstrated that although cytokinin was present in adventitious roots, storage root initiation only occurred when sucrose concentration was increased. These data suggest that transplants need to be fully established (SR1) and undergo initial canopy growth (SR2) before storage root initiation. Thus, it appears that any physical or biochemical variable that interferes with sucrose and cytokinin metabolism during SR1 and SR2 will lead to inconsistent storage root initiation, thereby influencing the final yield at a given harvest date. For example, partial root zone drying reduced zeatin and zeatin riboside in roots, shoot tips, and buds by 60%, 50%, and 70%, respectively, in grapevines (*Vitis vinifera*) (Stoll et al., 2000). Qiu et al. (2004) also demonstrated that chemical inputs such as pesticides altered zeatin riboside content in rice roots. Such information and further understanding of the molecular basis of storage root initiation will further assist in planning research studies that attempt to further identify the likely causes of unexpected or unexplained storage root variability or suboptimal yield at harvest.

The explicit recognition of the relationship between causal variables and the origin
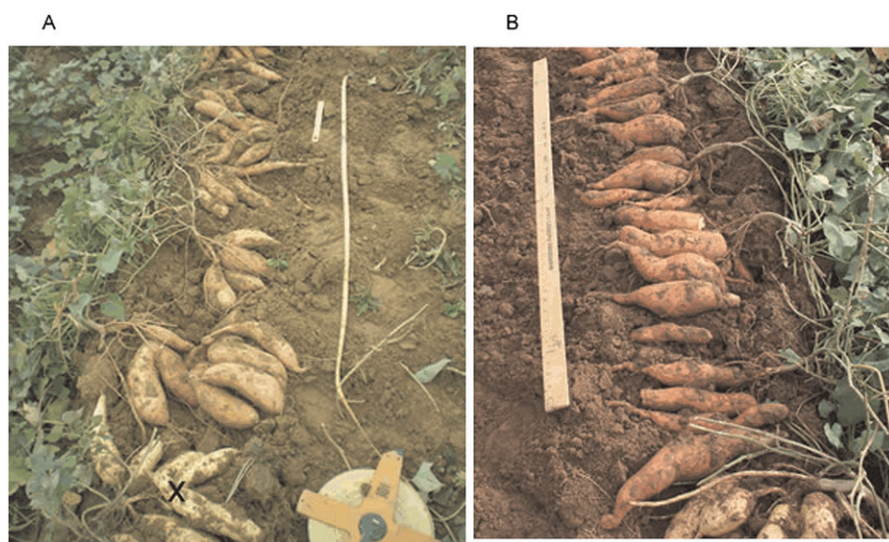

Fig. 7. Representative storage roots that were sampled at harvest from plots maintained near 50% of field capacity at 15-cm depth in Chase, LA. Samples from **A** were transplanted on 26 May 2009. Samples from **B** were transplanted on 15 June 2009. The hill marked "x" in **A** shows an 'O'Henry' guard plant. Agroclimatic and management variables are described in "Materials and Methods."

and early development of storage roots (phenology) underpinned the development of BxNET. These presumptive phenological stages have been locally validated and calibrated for the Beauregard cultivar. The prototype model is relatively simple, i.e., few total parameters, represents a limited range of agroclimatic conditions, and applies only to the specific management practices. This is consistent with the commonly held belief that the extrapolation potential of empirical models is considerably less than process models (Jame and Cutforth, 1996). As a beta-level model, BxNET needs to be further tested, calibrated, validated, and updated. Marcot et al. (2006) described two methods of updating BBN models with "case files," i.e., newly collected data: 1) using test results to calibrate model states to better align with the data; and 2) using case data to automatically update the CPTs. Further testing and validation should also be performed to determine model performance with different locations, planting densities, cultivars, and irrigation methods. BxNET can also serve as a foundation model for representing causal relationships of other variables such as various soil moisture regimes, weeds, disease, insects, and chemical injury. For example, there is a need to document yield response to prolonged soil moisture deficit (presumably below 8% to 10% VWC at 5- and 15-cm depths) and saturated conditions (presumably above 20% VWC at 5- and 15-cm depths) at SR1, SR2, SR3, and later growth stages. Togari (1950) documented the effects of marginal soil moisture 20 DAT and observed that "dryness and compactness of soil" increased cambium activity but led to increased lignification resulting in the formation of pencil-like storage roots. On the other hand, "shortage of O_2" led to decreased cambium activity and increased lignification, rendering adventitious roots prone to become non-storage roots (Togari, 1950). The expansion of BxNET to include other variables such as soil moisture can be accomplished by using the modular approach described by Kristensen and Rasmussen (2002). The current agroclimatic predictor variables can be grouped into a submodule. Subsequently, submodules representing soil moisture, fertility regimes, weed presence, disease incidence, etc., can be added as these become available. As the prototype BBN model is further developed and expanded, it can be converted into a decision network through the addition of decision and utility nodes. Utility nodes explicitly define values, e.g., net profit, associated with specific decision states. This capability represents an important characteristic of the BBN modeling paradigm: its ability to integrate agroclimatic, management, biological, and economic-themed variables into a comprehensive, systemwide model.

Crop growth models are important components of agriculture-related decision support systems (Park et al., 2005). Using BxNET to represent the likelihood of an outcome given certain scenarios helps to demonstrate the usefulness of BBNs in nat-ural resource management, i.e., depiction of the influence of alternative management interventions (e.g., altering harvest date) on key predictor variables, thereby enabling the manager to choose the best course of action (McCann et al., 2006). BBNs are not without drawbacks. One of the main disadvantages of using BBNs is the need to discretize continuous variables. Discretization could lead to lower precision (McCann et al., 2006). On the other hand, we gain the ability to use the efficient reasoning machinery of BBNs, especially if the relationships among variables are non-linear and complex (Myllymaki et al., 2002). Our data-driven approach of BBN development used a combination of findings from previous research and recent empirical data in helping to delimit the extent of data collection and variable selection. In the context of representing the response of a complex biological system, we believe that it was important to precisely define and confirm the onset of storage root initiation and to adopt a phenology-based modeling framework. Our results further corroborate Togari's (1950) findings that agroclimatic and management variables within the first 20 DAT directly influenced the rate of storage root initiation and eventual yield. We are able to make this comparison because of a common phenological timeframe. This approach helps to ensure that future calibration and validation work on the prototype model will be based on a common frame of reference.

BxNET and the accompanying modeling data set are available on request from the authors. A trial version of Netica can be downloaded from the vendor's web site. The software is needed to run and interact with BxNET.

Conclusion

We described the development of BxNET, a prototype or beta-level BBN model that represented the relationship between US1YIELD and agroclimatic variables extant during the critical storage root initiation stages. Environmental variability during the critical storage root initiation phase influenced storage root count and final storage root yield was a function of time, assuming other variables were uniform. The model was developed assuming the absence of influence from weeds, disease, insect pests, and chemical injury. BBN development was performed in two basic steps. First, current knowledge regarding the presumptive relationship between agroclimatic variables and storage root initiation was used to develop candidate alpha-level BBN structures. These candidate models were subsequently parameterized on a training data set and a beta-level model was identified using a search and score approach. Model error rates were estimated using a combination of cross-validation (leave-one-out), validation on an independent data set, and AUC analysis. As an empirically derived model, BxNET is applicable only to the location defined in the study. As a beta-level model, BxNET needs to undergo further testing, calibration, validation, and updating. Site-specific calibration and validation will allow BxNET to be adopted for use in other locations or growing environments. The model can be used as a foundation for investigating the influence of other variables such as weed presence, disease incidence, and other yield-limiting factors. It can also be used as a basis for the development of a model-based decision support system to further increase efficiency in sweetpotato production.

Literature Cited

Binder, J., D. Koller, S. Russell, and K. Kanazawa. 1997. Adaptive probabilistic networks with hidden variables. Mach. Learn. 29:213–244.

Bouwkamp, J. 1985. Production requirements, p. 1–33. In: Bouwkamp, J. (ed.). Sweetpotato products: A natural resource for the tropics. CRC Press, Inc., Boca Raton, FL.

Cain, J. 2001. Planning improvements in natural resources management: Guidelines for using Bayesian networks to support the planning and management of development programmes in the water sector and beyond. Centre for Ecology and Hydrology. Crowmarsh Gifford, Wallingford, UK.

Connell, J.L. and L.B. Shafer. 1989. Structured rapid prototyping. Prentice-Hall, Englewood Cliffs, NJ.

Constantin, R.J., T.P. Hernandez, and L.G. Jones. 1974. Effects of irrigation and nitrogen fertilization on quality of sweet potatoes. J. Amer. Soc. Hort. Sci. 99:308–310.

Dufault, R.J. 1997. Determining heat unit requirements for broccoli harvest in coastal South Carolina. J. Amer. Soc. Hort. Sci. 122:169–174.

Eguchi, T., M. Kitano, and H. Eguchi. 1994. Effect of root temperature on sink strength of tuberous root in sweet potato plants (*Ipomoea batatas* Lam.). Biotronics 23:75–80.

Eguchi, T., M. Kitano, and H. Eguchi. 1998. Growth of sweetpotato tuber as affected by the ambient humidity. Biotronics 27:93–96.

Eguchi, T. and S. Yoshida. 2008. Effects of application of sucrose and cytokinin to roots on the formation of tuberous roots in sweetpotato (*Ipomoea batatas* L. Lam.). Plant Root 2:7–13.

Gu, Y., D.R. Peiris, J.W. Crawford, J.W. McNichol, B. Marshall, and R.A. Jefferies. 1994. An application of belief networks to future crop production, p. 305–309. Proc. 10th Conf. on. Artificial Intelligence for Applications, San Antonio, TX.

Jame, Y. and H.W. Cutforth. 1996. Crop growth models for decision support systems. Can. J. Plant Sci. 76:9–19.

Jenni, S., D. Cloutier, G. Bourgeois, and K.A. Stewart. 1996. A heat unit model to predict growth and development of muskmelon to anthesis of perfect flowers. J. Amer. Soc. Hort. Sci. 2:274–280.

Ji, B., Y. Sun, S. Yang, and J. Wan. 2007. Artificial neural networks for rice yield prediction in mountainous regions. J. Agr. Sci. 145:249–261.

Kays, S.J. 1985. The physiology of yield in the sweetpotato, p. 79–132. In: Bouwkamp, J. (ed.). Sweetpotato products: A natural resource for the tropics. CRC Press, Boca Raton, FL.

Kristensen, K. and I.A. Rasmussen. 2002. The use of a Bayesian network in the design of a decision support system for growing malting barley without use of pesticides. Comp. Elec. Agr. 33:197–217.

Kuhnert, P.M. and K.R. Hayes. 2009. How believable is your BBN?, p. 4319–4325. In: Anderssen, R.S., R.D. Braddock, and L.T.H. Newham (eds.). 18th World IMACS Congress and MODSIM09 Int. Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, July 2009, Cairns, Australia.

Marcot, B.G., J.D. Steventon, G.D. Sutherland, and R.K. McCann. 2006. Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. Can. J. For. Res. 36:3063–3074.

McCann, R.K., B.G. Marcot, and R. Ellis. 2006. Bayesian belief networks: Applications in ecology and natural resource management. Can. J. For. Res. 36:3053–3062.

Mcharo, D. and D. LaBonte. 2007. Genotypic variation among sweetpotato clones for B-carotene and sugar content, p. 764–754. In: Kapinga, R., R. Kingamkono, M. Msabaha, J. Ndunguru, B. Lemaga, and G. Tusiime (eds.). Proc. 13th Int. Soc. Trop. Root Tuber Crops, 10–14 Nov. 2003, Arusha, Tanzania.

Mills, H.A. and J.B. Jones, Jr. 1996. Plant analysis handbook II. Practical sampling, preparation, analysis, and interpretation guide. Micro-Macro Publishing, Athens, GA.

Mithra, V.S.S. and K. Somasundaram. 2008. A model to simulate sweet potato growth. World Appl. Sci. J. 4:568–577.

Myllymaki, P., T. Silander, H. Tirri, and P. Uronen. 2002. B-Course: A web-based tool for Bayesian and causal analysis. Int. J. Artif. Intell. Tools 11:369–387.

NASS. 2009. National Agricultural Statistics Service. 20 Jan. 2010. <http://www.nass.usda.gov/>.

Neopolitan, R.E. 2003. Learning Bayesian networks. Prentice Hall, Englewood Cliffs, NJ.

Park, S.J., C.S. Hwang, and P.L.G. Vlek. 2005. Comparison of adaptive techniques to predict crop yield response under varying soil and land management conditions. Agr. Syst. 85:59–81.

Qiu, Z., J. Wu, B. Dong, D. Li, and H. Gu. 2004. Two-way effect of pesticides on zeatin riboside content in both rice leaves and roots. Crop Prot. 23:1131–1136.

Ravi, V. and P. Indira. 1999. Crop physiology of sweet potato, p. 277–316. In: Janick, J. (ed.). Horticultural reviews. Vol. 23. John Wiley & Sons, Inc., New York, NY.

Rolston, L.H., C.A. Clark, J.M. Cannon, W.M. Randle, E.G. Riley, P.W. Wilson, and M.L. Robbins. 1987. 'Beauregard' sweet potato. HortScience 22:1338–1339.

Singh, U., J.T. Ritchie, and G. Tsuji. 1992. Simulation models for crop growth: The IBSNAT approach, p. 343–355. In: Hill, W.A., C.K. Bonsi, and P.A. Loretan (eds.). Sweetpotato technology for the 21st century. Tuskegee University, Tuskegee, AL.

Somasundaram, K. and V.S. Mithra. 2008. Madhuram: A simulation model for sweet potato growth. World J. Ag. Sci. 4:241–254.

Stoll, M., B. Loveys, and P. Dry. 2000. Hormonal changes induced by partial rootzone drying of irrigated grapevine. J. Expt. Bot. 51:1627–1634.

Tanaka, M., N. Kato, H. Nakayama, M. Nakatani, and Y. Takahata. 2008. Expression of class I knotted1-like homeobox genes in the storage roots of sweetpotato (*Ipomoea batatas*). J. Plant Physiol. 165:1726–1735.

Tari, F. 1996. A Bayesian network for predicting yield response of winter wheat to fungicide programmes. Comp. Elec. Agr. 15:111–121.

Togari, Y. 1950. A study of tuberous root formation in sweet potato. Bul. Nat. Agr. Expt. Sta. Tokyo 68:1–96.

Turban, E. 1992. Expert systems and applied artificial intelligence. MacMillan, New York, NY.

U.S. Department of Agriculture. 2005. United States standards for grades of sweetpotatoes. 18 Aug. 2008. <http://www.ams.usda.gov/AMSv1.0/getfile?dDocName=STELPRDC5050330>.

Uusitalo, L. 2007. Advantages and challenges of Bayesian networks in environmental modelling. Ecol. Model. 203:312–318.

Villordon, A., J.M. Cannon, H.L. Carroll, J.W. Franklin, C.A. Clark, and D.R. LaBonte. 2003. Sweetpotato 'Beauregard' mericlones vary in yield, vine characteristics, and storage root size and shape attributes. HortScience 38:1089–1092.

Villordon, A., D.R. LaBonte, and N. Firon. 2009a. Development of a simple thermal time method for describing the onset of morpho-anatomical features related to sweetpotato storage root formation. Sci. Hort. 121:374–377.

Villordon, A., D.R. LaBonte, N. Firon, Y. Kfir, E. Pressman, and A. Schwartz. 2009b. Characterization of adventitious root development in sweetpotato. HortScience 44:651–655.

Villordon, A., C. Clark, D. Ferrin, and D. LaBonte. 2009c. Using growing degree days, climatic variables, linear regression, and data mining methods to help improve prediction of sweetpotato harvest date and yield in Louisiana. HortTechnology 19:133–144.

Watkins, M.W. 2000. An Excel program for calculating and graphing the receiver operating characteristic (ROC). Ed & Psych Associates, State College, PA.

Wilson, L.A. and S.B. Lowe. 1973. The anatomy of the root system in West Indian sweet potato [*Ipomoea batatas* (L.) Lam] root system during early growth from stem cuttings. Ann. Bot. (Lond.) 37:633–643.

Zhang, N.L., T.D. Nielsen, and F.V. Jensen. 2004. Latent variable discovery in classification models. Artif. Intell. Med. 30:283–299.