Atlantic Electronic
Journal of Mathematics
Volume **5**, Number **1**, Winter **2012**

http://aejm.ca
http://aejm.ca/rema
pp. **77–89**

SIMPLE STOCHASTIC MODELS OF FOOD INSECURITY DUE TO CLIMATE CHANGE

David Ddumba Walakira

Department of Mathematics
Makarere University
P.O. Box 7062
Kampala, Uganda

Andrew Kayanja

Department of Mathematics
Gulu University
P.O. Box 166
Gulu, Uganda

Marek Stastna

Department of Applied Mathematics
University of Waterloo
200 University Avenue West
Waterloo, ON, N2L 3G1

Abstract. While climate change is a complex, interdisciplinary topic, perhaps the most important effect of climate change on the developing world is through the issue of food security. We present some simple mathematical models for climate-change-induced food insecurity in the equatorial region. It is believed that the primary effect of future climate change in equatorial Africa will be a shift of the timing of the rainy season and an increase in the incidence of violent rain events. By considering a stochastic parameter dependence in simple models of crop yield we demonstrate the manner in which the distribution of crop yield varies. In particular we demonstrate that in some parameter regimes an unexpected double-peaked distribution results, with a significant probability of potentially disastrous low yields. We conclude by critiquing our simple models and suggesting avenues for future work.

1. **Introduction.** The issue of climate change is the exemplar of interdisciplinarity, spanning the physical, biological and computer sciences along with a profound political dimension. The fourth IPCC report (Solomon *et al. 2007*) identifies several future climate scenarios. In all but the most optimistic of these, the world's climate is expected to become far less regular, with the heavily populated regions of tropical and subtropical Africa experiencing profound and difficult-to-predict changes due to the shifts in atmospheric convergence bands. This is particularly problematic since many developing economies are not well equipped to deal with resulting crop failures, water shortages and a host of other issues lumped under the heading of "food security".

2000 *Mathematics Subject Classification.* Primary: 35Q51, 76B15; Secondary: 76U05.
Key words and phrases. food security, stochastic models, climate change.

The modeling of crops, while not as well known as physics-based computer models of the climate system, is a vibrant part of so-called land surface models (Chalinor *et al* 2004), and is closely tied to the manner in which meteorological modeling efforts (Slingo *et al* 2005) are used to infer implications for agriculture, the spread of disease and other so-called human factors. Much as their climate system counterparts, crop models exhibit varying degrees of nonlinearity and number of parameters. However even so-called intermediate complexity models (e.g. Chalinor *et al* 2004) require well over twenty parameters, many of which are only weakly constrained. The model of Chalinor *et al*, is specifically designed to minimize parameter calibration, while more complex models (e.g. CROPGRO, Boote and Jones 1998) sometimes have entire calibration methodologies (Basso *et al* 2001) in order to specify the model parameters in a robust manner

In this article we develop several very simple models of crop yield using standard equations and methods from population and mathematical biology. The focus is on the simplest models, building on long accepted practice in climate modeling (e.g. Stommel's two box model, Stommel 1961, the delayed oscillator model of the El Nino-Southern Oscillation, ENSO, Tziperman *et al* 1994). No attempt is made to incorporate modern farming practices, including genetically modified seeds and fertilizer application regimes, which, in any event, are largely out of reach for typical farmers in developing countries. The role of climate change is introduced via stochastic parameter dependencies and a variety of analytical and semi-analytical results are derived. We identify mathematical issues that preclude a completely analytical treatment, and at the same time, demonstrate the manner in which many of our results carry over to wider classes of models. We conclude by criticizing our models, and based on this criticism, identify avenues for future work.

2. Models.

2.1. The simplest model.

Consider the total crop yield (biomass) for a typical small plot of land in the tropics, tilled either by hand or with simple machinery. Assume no fertilizers or pesticides are applied and that the biomass in the field increases exponentially depending on the availability of moisture. Further assume that at the time of planting $t = 0$ there is little rain, and a small biomass Y_0. Steady rains set in at a time $0 < t_1 < 1$ and continue until harvest at $t = 1$. We have thus scaled time, and biomass so that our resulting models will be dimensionless. For $t < t_1$ we have

$$\frac{dY}{dt} = \alpha_1 Y$$

so that $Y(t_1) = Y_0 \exp(\alpha_1 t_1)$ and for $t > t_1$ we have

$$\frac{dY}{dt} = \alpha_2 Y$$

where $\alpha_2 \gg \alpha_1$ so that at $t = 1$ we have

$$Y(1) = Y_0 \exp(\alpha_1 t_1) \exp(\alpha_2 [1 - t_1]) = Y_0 \exp[-t_1(\alpha_2 - \alpha_1) + \alpha_2].$$

Let us now consider a probability distribution for the time of onset of the rains, t_1, with a particularly simple choice being $t_1 \sim N(1/2, \sigma^2)$, a Gaussian distribution centered at 1/2 with a standard deviation of σ. What distribution does $Y(1)$ have? Notice that we can rewrite the solution as

$$\ln\left(\frac{Y(1)}{Y_0}\right) = \alpha_2 - t_1(\alpha_2 - \alpha_1)$$

where the right hand side has a Gaussian distribution $N(\alpha_2 - (\alpha_2 - \alpha_1)/2, (\alpha_2 - \alpha_1)\sigma^2)$, or in other words $Y(1)/Y_0$ has a lognormal distribution. The expression may be simplified further if we assume that $\alpha_2 = \beta\alpha_1$ and we scale by the value one would get were the crop to grow exponentially under the optimal growth conditions. This means our scaled variable will take a value between 0 and 1, so that its logarithm is negative, and the probability distribution is given by

$$\ln\left(\frac{Y(1)}{Y_0\exp(\alpha_1\beta)}\right) \sim N(\alpha_1(1-\beta)/2, \sigma^2\alpha_1(\beta-1)).$$

Note that the right hand has a standard deviation that is positive (since $\beta > 1$) and is directly proportional to the standard deviation of the onset of rains (σ), the pre-rain growth parameter α_1 and the post-rain growth parameter $\beta\alpha_1$. Several examples are presented in figure 1. These clearly demonstrate that for large differences between the optimal ($\beta\alpha_1$) and sub-optimal (α_1) growth rates the probability of a highly sub-optimal yield increases very quickly, so that even moderate variance in the timing of the rains has possibly grave consequences.

2.2. **Logistic model.** The simple model presented above can be criticized in many different ways. Perhaps the most obvious criticism is that the crop biomass cannot increase without bound and indeed every field has some inherent carrying capacity that provides a hard upper bound on crop yield. It is well known that the logistic model is the simplest extension of the exponential model with a carrying capacity (Murray 2002). Since a logistic model reduces to an exponential model when biomass Y is small, we only modify the post-onset of the rains growth law from our simplest model, so that for $t > t_1$

$$\frac{dY}{dt} = \beta\alpha_1 Y(1-Y).$$

which can be solved to yield

$$Y(1) = -\frac{e^{\alpha_1 t_1} y_0 e^{-\alpha_1\beta\, t_1}}{-e^{\alpha_1 t_1} y_0 e^{-\alpha_1\beta\, t_1} - e^{-\alpha_1\beta} + e^{-\alpha_1\beta} e^{\alpha_1 t_1} y_0}.$$

We have not been able to express the distribution of $Y(1)$ in terms of standard probability distributions, even when t_1 has a Gaussian distribution. A formula can be derived by the change-of-variable theorem, or samples can be generated in standard mathematical software packages such as Maple and Matlab. Both of these have been used to generate figures in the following. An ensemble size of 9×10^6 samples was used to obtain robust statistics throughout the following.

Since growth now has a hard bound we expect that parameter sets with larger growth rates, which are more likely to run up against the hard limit on Y, will yield qualitatively different probability distributions. This is confirmed in figure 2. When β is small, the probability distribution is peaked at low values. This is because whether rains come early or late, the biomass grows slowly. However, when the growth rate is increased, a far broader probability distribution results, and for larger growth rates still, the probability distribution peaks at the carrying capacity.

2.3. **Modified Logistic model.** The above-described logistic model is an obvious improvement on the purely exponential model. Nevertheless, it ignores several key aspects of the anticipated climate change assumed to affect the modeled crop yield. In particular, delayed rains can lead to the death of a subset of the planted crop (e.g. seeds, seedlings, juvenile plants), and (Solomon *et al. 2007*) delayed rains are often accompanied by higher incidence of extreme weather events, which can in turn wash away a portion of the planted crop. The simplest model of both of these effects is a carrying capacity that decreases as t_1 increases above its mean value, so that

$$\frac{dY}{dt} = \beta\alpha_1 Y(a - Y).$$

where

$$a(t_1) = 1 - H(t_1 - 1/2)b[t_1 - 1/2]$$

and $H(.)$ is the Heaviside step function.

The empirical function $a(t_1)$, which specifies the carrying capacity, is chosen so that early rains have no effect. The strength of influence of the later onset of rains is controlled by the single parameter b. The linear nature of $a(t_1)$ is the simplest possible choice. From the form of the model, we can predict that the most profound differences with the pure logistic model can be expected when the growth rate is large and the decrease in the carrying capacity is significant, $b > 1$. In these cases we can expect an increase in the probability of low yields. This is confirmed in figure 3, where we contrast two modified model parameter sets with the purely logistic model. It is clear that for the largest values of b the probability distribution has a second peak that is centered very near zero. Even for significantly reduced values of b (dashed curve) the probability distribution is bimodal, with a peak near optimal yields, a very sharp drop off near $Y(1) \approx 0.7$ and a secondary peak centered well below $Y(1) = 0.1$. By examining the individual ensemble members, we have confirmed that the high yield peak corresponds to ensemble members for which the rains set in early $t_1 < 1/2$.

In figures 4 and 5 we show isolines of probability that $Y(1) < 0.1$ and $Y(1) < 0.5$, respectively. These clearly show that the modified logistic model only leads to significant differences from the purely logistic model when the growth parameter β is large enough. Moreover, the effects on the probability of very low yields $Y(1) < 0.1$ cover a much larger section of parameter space.

2.4. **Modifying the Probability distribution of the rains.** While the Gaussian distribution for the onset of the rains is the simplest it has several undesirable properties: it is symmetric and has so-called fat tails, or instances in which draws from it result in extreme values. One way to draw from a fixed interval is to use the so-called Beta distribution. For the Beta distribution in standard form the probability distribution function is zero outside the interval $[0, 1]$ and is given by

$$P(t; \nu, \omega) = \frac{t^{-1+\nu}\,(1-t)^{-1+\omega}}{\beta\,(\nu, \omega)},$$

where $\beta(\nu, \omega)$ is a constant given in terms of the Gamma function by

$$\beta(\nu, \omega) = \frac{\Gamma(\nu)\Gamma(\omega)}{\Gamma(\nu + \omega)}.$$

The parameters ν and ω determine the shape of the probability distribution and the range $[0, 1]$ can be linearly shifted to any interval of choice ($0.475 < t_1 < 0.775$ for the figure).

In figure 6 we show results for the three previously discussed models with the beta distribution used in place of the normal. In panel (a) we show the histogram for the 9×10^6 draws from the beta distribution used to construct subsequent panels. We set $(\nu, \omega) = (2, 3.2)$, though similar results would be obtained for many different parameter values. In panel (b) we show results from the exponential model. In contrast to figure 1 we find that even low growth rates yield sharply peaked probability distributions for yield. In panel (c) we show results for the purely logistic model. It can be seen that again the probability distributions are more peaked than those for the Gaussian t_1, especially at low growth rates. This is consistent with results shown in panel (b). In panel (d) we show results for the modified logistic model. It can be seen that over a substantial range of b values a double-peaked probability distribution results. Interestingly the peak for large yields is substantially reduced in extent from results shown in figure 3. A direct examination of the t_1 values that give yields larger than 0.85 show that these are never larger than 0.51. This implies that the broad peak of large yields in figure 3 is a direct result of the symmetric nature of the Gaussian, and is probably unrealistic.

2.5. **The effect of a nonlinear carrying capacity.** Previously we have considered the decrease in carrying capacity to be a linear function in the delay of the onset of the rains $(t_1 - 1/2)$. This is the simplest model, though unlikely to be realistic. While many possible models for $a(t_1)$ are possible, one in which a decreases quickly once a critical delay is reached is the logical counterpoint to the linear model discussed above. We thus consider

$$\frac{dY}{dt} = \beta\alpha_1 Y(a - Y),$$

where

$$a(t_1) = (1 - B) + B \tanh\left(\frac{t_1 - b}{d}\right).$$

Since the hyperbolic tangent varies between -1 and 1, the parameter $2B$ determines the size of the decrease in a, the parameter b determines the t_1 value at the center of the change, and d determines how sharply a decreases. We wish to consider large, rapid variations in a and hence consider $B = 0.45$ and $d = 0.05$ and allow b to vary. The onset of the rains is taken as being drawn from the same beta distribution as the results shown in figure 7. In panel 7 a we show the probability distribution of $Y(1)$ for various choices of b. The results should be compared those in figure 6d. It can be seen that the primary effect of the highly nonlinear change in a is to decrease the high yield peak, and to decrease the peak for very low yields. This is especially true as b decreases. Panel 7b shows the low yield range in detail while panel 7c shows $a(t_1)$. As the width of the hyperbolic tangent in the specification of a is increased the linear results from figure 6 may be recovered (not shown).

3. **Discussion.** The above discussed models clearly demonstrate that the uncertainty in the onset of and increase in violent rain events during the rainy season (the so-called long rains in the equatorial East African context) yield parameter regimes for which the probability of catastrophically low yields is greatly increased. The models constructed exhibit this property due to the variation in the maximum attainable yield due to, for example, seed death due to drought and seed washout due to flooding. Due to their simplicity, our models do not require parameter fitting, and indeed are not designed to fit any particular data set. Indeed, detailed predictions are best left to well calibrated models that form a part of a modern meteorological observation, data assimilation and prediction network. However, many developing countries (e.g. Uganda) lack the infrastructure and resources for such a network, and for such settings simple models can convey information to a variety of stakeholders in a compact manner.

While the presented models should eventually be superceded by a multi-stake holder modeling effort driven by the needs of the local, national entities, the present work could be extended in a number of useful ways that do not require vast expenditures or infrastructure. The primary extension to a tropical setting is via consideration of different types of crops. For example, while subsistence farming of maize can be reasonably be considered to be roughly reflected in the models presented above, more economically complex crops such as coffee and bananas cannot be. The construction of simple models that link the economics of cash crops with "cartoons" of predicted climate change form the most immediate extension of the present work.

On the climate side it would be worth comparing the stochastic perturbations due to either observed or modelled ENSO system variability, including toy models such as the delayed oscillator (Tziperman *et al* 1994), to the idealized perturbations considered above. Moreover, due to the well known problems many climate models have with the correct characterization of the Inter-Tropical Convergence Zone (the so-called double ITCZ problem, see Wu *et al* 2007) a serious nested modeling effort focusing on Tropical Africa is needed.

Acknowledgements. The authors would like to acknowledge the support of the African Initiative of the Centre for International Governance Innovation (CIGI) for (DDW and AK) and NSERC (M.S.).

REFERENCES

[1] Basso, B., Ritchie, J.T., Pierce, F.J., Braga, R.P., Jones, J.W., *Spatial validation of crop models for precision agriculture*, Agricultural Systems, **68**, 2001, 97-112.

[2] Boote, K.J. & Jones, J.W. Simulation of crop growth: CROPGRO model. in *Agricultural system modeling and simulation*, ed. Peart, R.M., Curry, R.B., Marcel-Dekker, New York, 1998.

[3] Chalinor, A.J., Wheeler, T.R., Craufurd, P.Q., Slingo, J.M., D.I.F. Grimes, *Design and optimisation of a large-area process-based model for annual crops*, Agricultural and Forest Meteorology, **124**, 2004, 99-120

[4] Murray, J.D., *Mathematical Biology I. An Introduction*, (2001), Springer, New York.

[5] Slingo, J.M., Challinor, A.J., Hoskins, B.J., Wheeler, T.R., *Introduction: food crops in a changing climate*, Phil. Trans. R. Soc. B, **360**, 2005, 1983-1989.

[6] Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M. & Miller, H.L., eds. *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental panel on Climate Change 2007*, Cambridge University Press, (2007), Cambridge, U.K.

[7] Stommel, H. *Thermohaline convection with two stable regimes of flow.* Tellus, **13**, (1961), 224-230.

[8] Tziperman, E., Stone, L., Cane, M.A., Jarosh, H., *El Niño Chaos: Overlapping of resonances between the seasonal cycle and the Pacific ocean-atmosphere oscillator*, Science, **264**, (1994) 72-74.

[9] Wu, X., Deng, L., Song, D., Vettoretti, G., Peltier, W.R., Zhang, G.J., *Impact of a modified convective scheme on the Madden-Julian Oscillation and El NIno-Southern Oscillation in a coupled climate model*, Geophys. Res. Lett., **34**, L16823, doi:10.1029/2007GL030637.

E-mail address: mmstastn@uwaterloo.ca; kayanja.andrew@yahoo.com, davidddumba@gmail.com

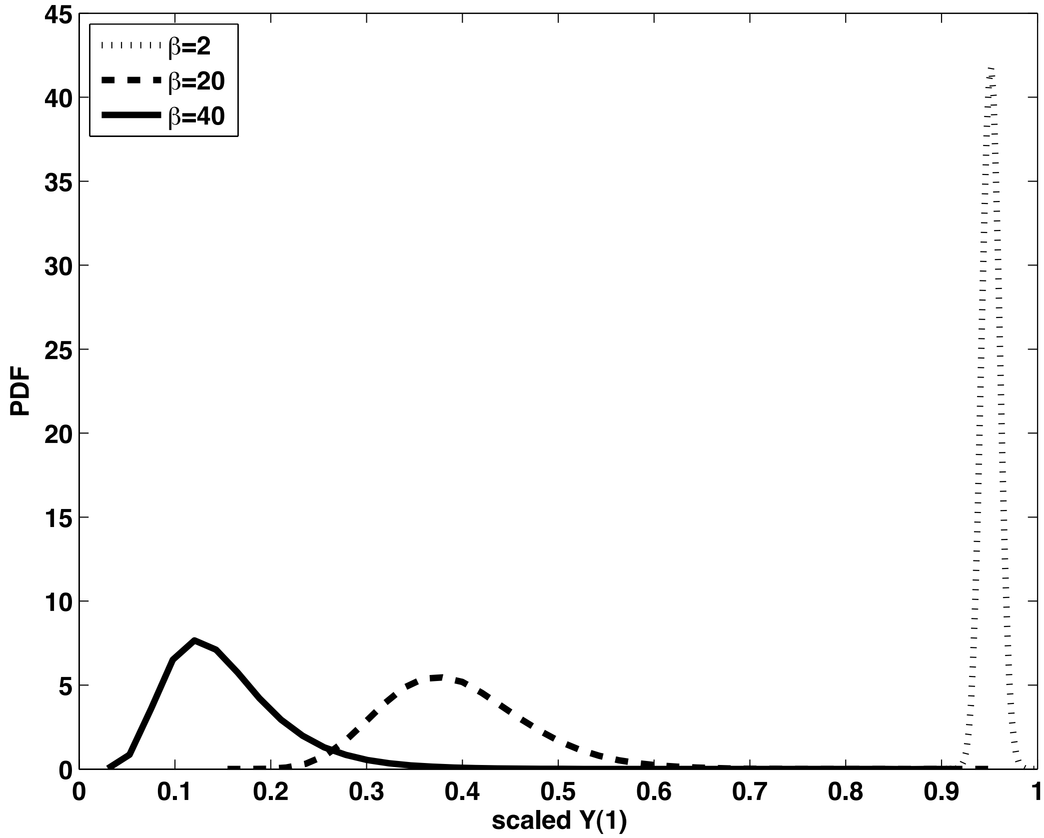

FIGURE 1. Sample PDFs for $Y(1)/Y(0)\exp(\beta\alpha_1)$ from the simplest model as β varies. $\alpha_1 = 0.1$ and $\sigma = 0.1$. Note how the effects of delayed rains increase for larger disparities between the sub-optimal (α_1) and optimal ($\beta\alpha_1$) growth parameters.

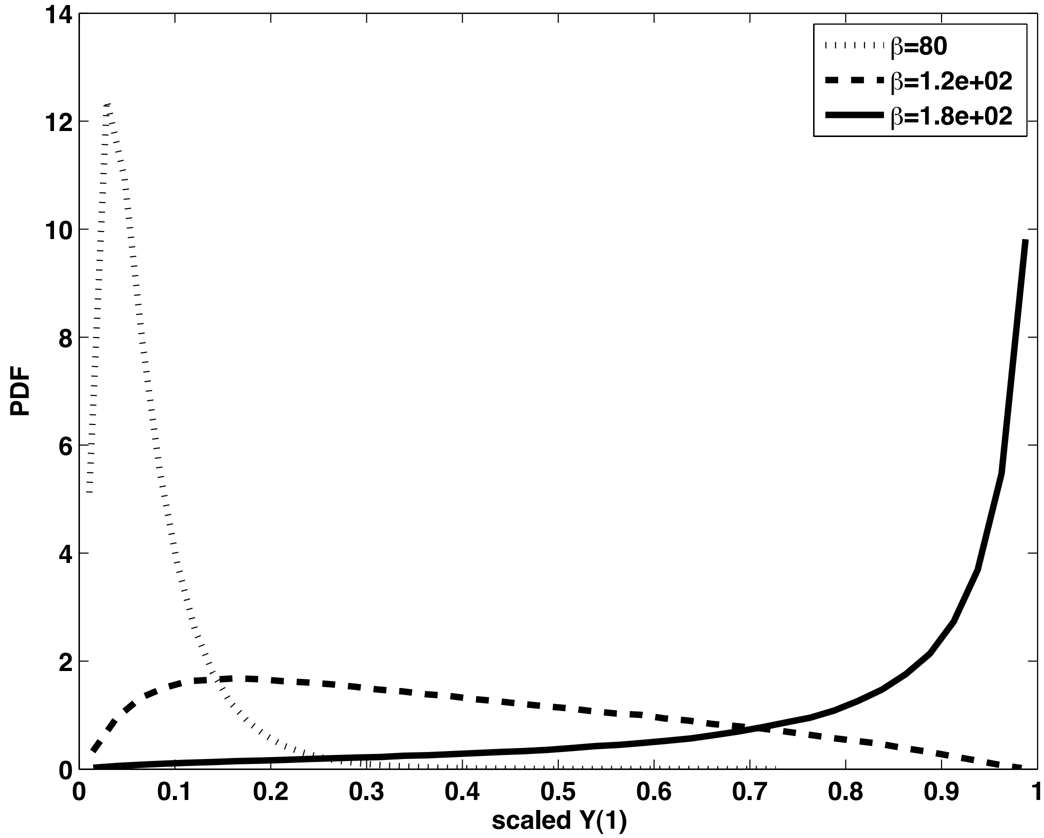

FIGURE 2. Sample PDFs for $Y(1)$ from the logistic model as β varies. $\alpha_1 = 0.1$ and $\sigma = 0.1$. Note how the effects of delayed rains can be qualitatively different depending on the growth rate.

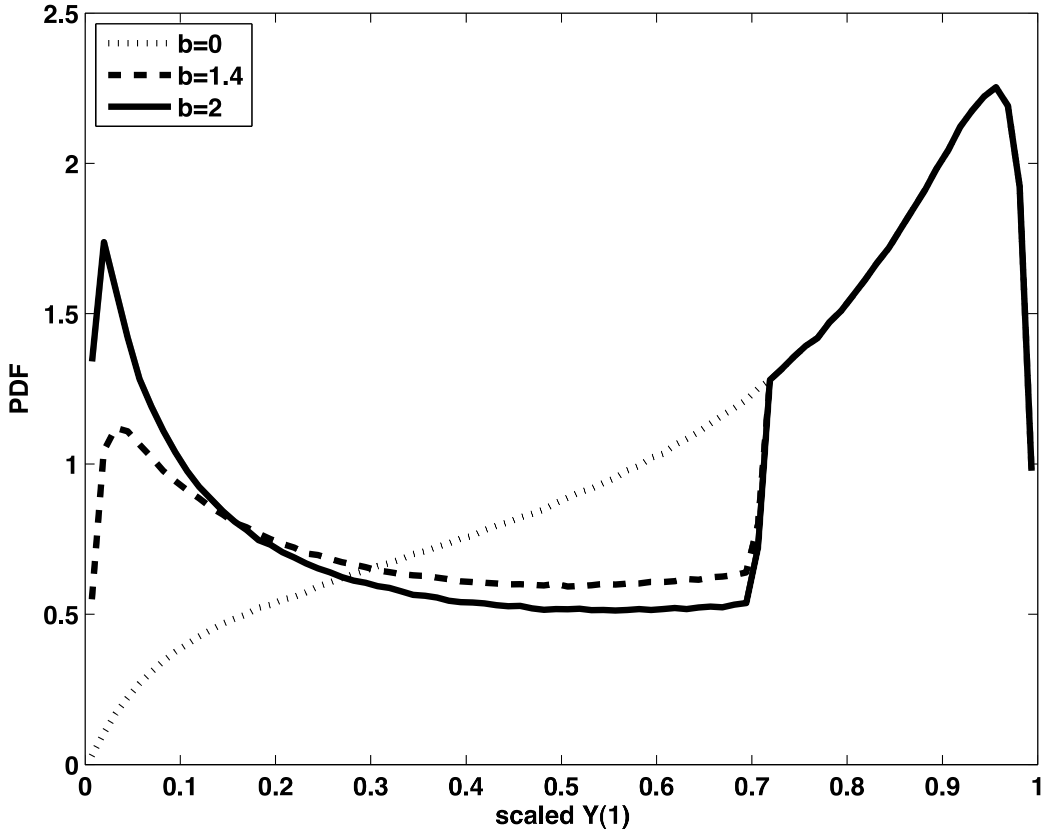

FIGURE 3. Sample PDFs for $Y(1)$ from the modified logistic model as b varies. $(\beta, \alpha_1) = (155, 0.1)$ and $\sigma = 0.1$. The dotted line correspond to the pure logistic model of figure 2. Note how the effect of delayed rains on the carrying capacity can lead to a double peaked probability distribution.

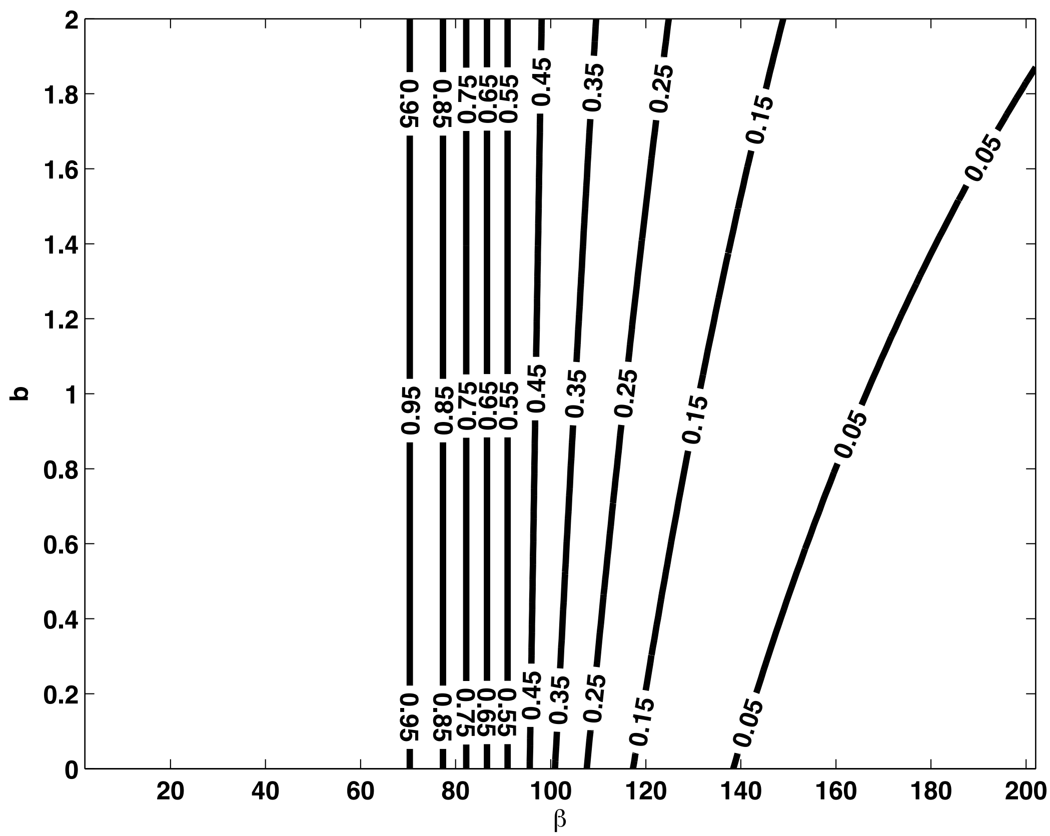

FIGURE 4. Contours of probability that $Y(1) < 0.1$ in $\beta - b$ parameter space. Note how the effect of delayed rains on the carrying capacity is only evident for large enough β.

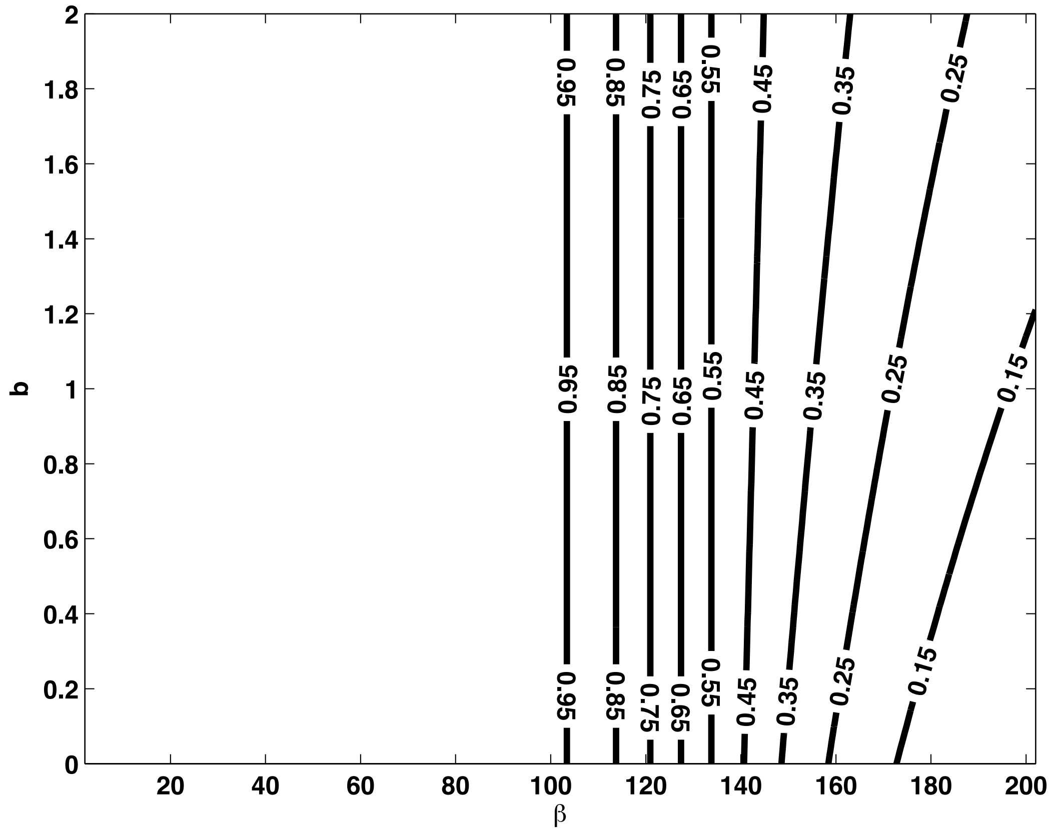

FIGURE 5. Contours of probability that $Y(1) < 0.3$ in $\beta - b$ parameter space. Note how the effect of delayed rains on the carrying capacity is only evident for the largest values of β (contrast with figure 4).

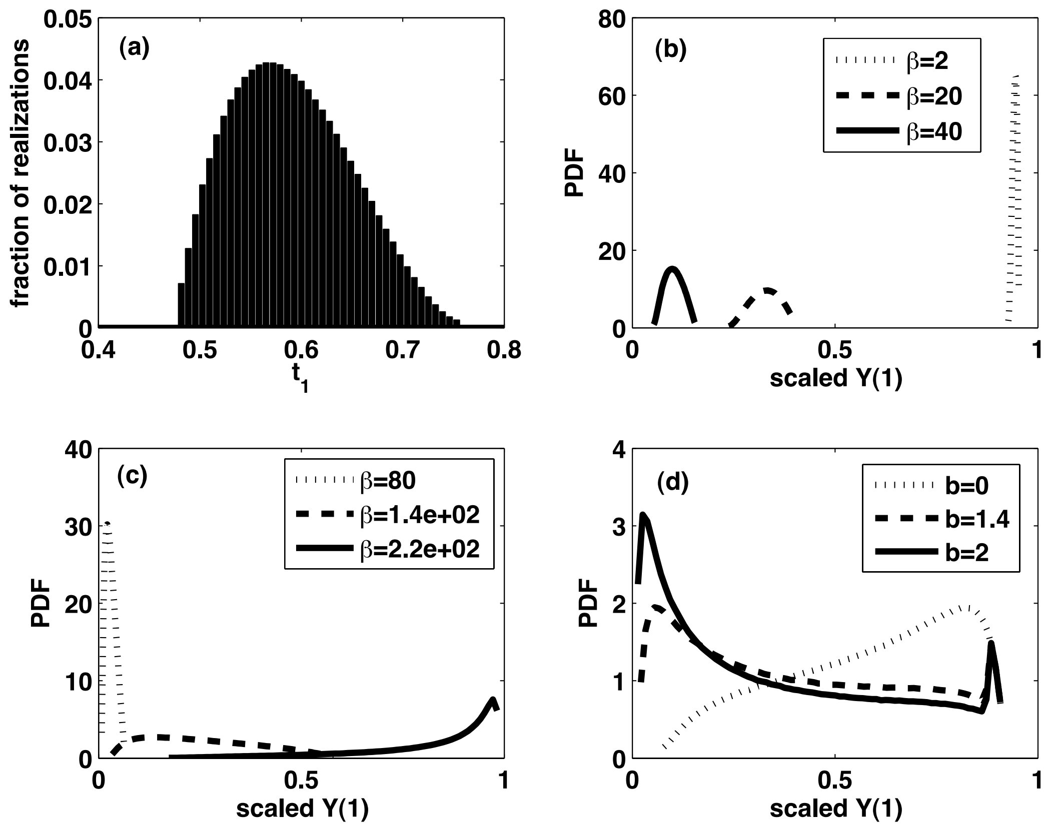

FIGURE 6. (a) Histogram of realizations from the beta distribution used in place of a Gaussian, (b) results of exponential model; contrast with figure 1, (c) results of purely logistic model with $\beta = 176$, contrast with figure 2, (d) results of modified logistic model; contrast with figure 3.

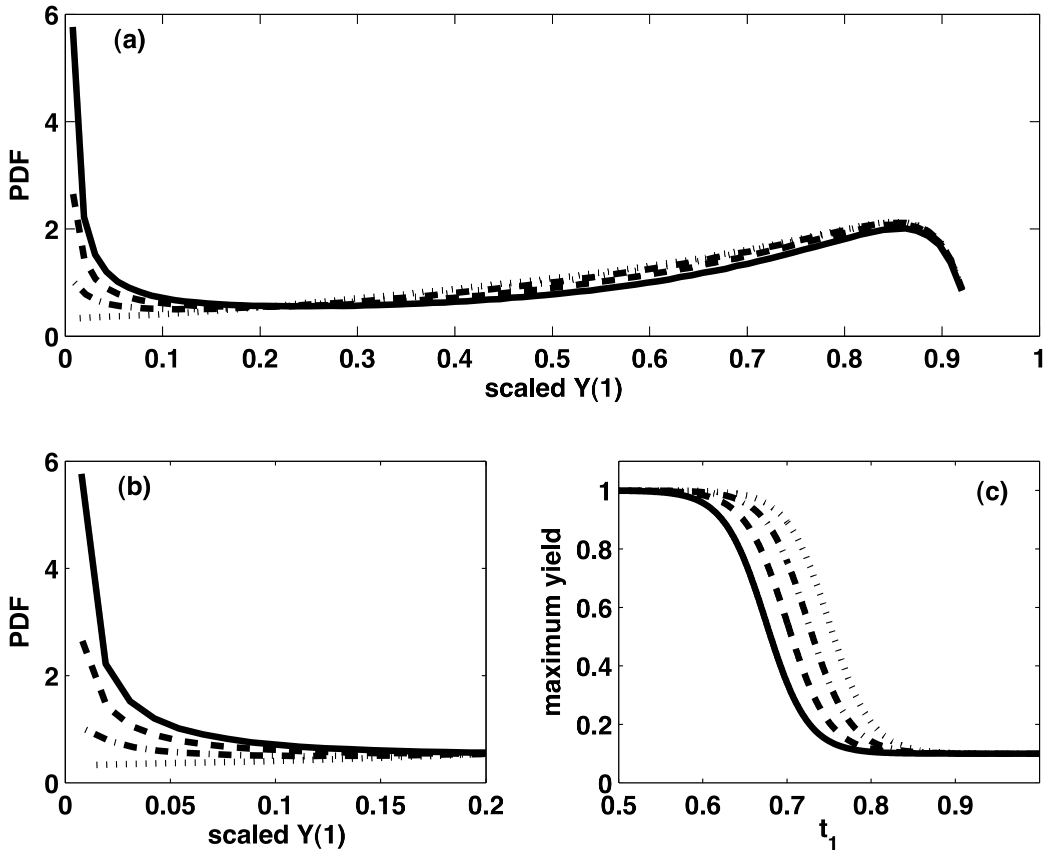

FIGURE 7. Results of a modified logistic model with a highly nonlinear change in maximum yield for fixed beta distribution and various choices of yield function. (a) Sample PDFs for $Y(1)$ from the modified logistic model as the center of a hyperbolic tangent b function varies, (b) detail of (a) for low yields, (c) $a(t_1)$, linestyles match profiles in panels (a) and (b).