HarvestChoice
BETTER CHOICES, BETTER LIVES

TECHNICAL NOTE

Measuring the Global Occurrence and Probabilistic Consequences of Wheat Stem Rust

April 10, 2013

Jason M. Beddow,[1,2] Terry M. Hurley,[1,2] Darren J. Kriticos,[1,2,3] and Philip G. Pardey,[1,2,4,5,6*]

Authorship is alphabetical. [1]International Science and Technology Practice and Policy (InSTePP) Center in the Department of Applied Economics at the University of Minnesota; [2]Commonwealth Scientific and Industrial Research Organization (CSIRO): [3]E.H. Graham Center at Charles Sturt University at Wagga Wagga: [4]Stakman-Bourlaug Cereal Rust Center at the University of Minnesota:[5] Department of Agricultural Economics, Extension and Rural Development, University of Pretoria; and [6] Department of Economics, University of Adelaide

The authors thank Marty Carson, Etienne Duveiller, David Hodson, Robert Park, Ravi Singh, Brian Steffenson, Bob Sutherst and participants at the American Phytopathology Society's Field Crops Rust Symposium in December 2011, San Antonio, Texas for their helpful insights, and Connie Chan-Kang, Michelle Hallaway and Noboru Ota for their excellent research assistance. This paper was prepared for the HarvestChoice project with support from the Bill and Melinda Gates Foundation, the University of Minnesota and the Commonwealth Scientific and Industrial Research Organization (CSIRO).

Correspondence should be addressed to beddow@umn.edu.

Measuring the Global Occurrence and Probabilistic Consequences of Wheat Stem Rust

This technical note describes the methods and data sources used to generate the results reported in Pardey et al. (2013). It includes details on a) the climate suitability model for *Puccinia graminis,* the causal agent for wheat stem rust[1], b) the source of geo-referenced wheat area, production and yield data used in this study, c) the methods we used to develop probabilistic estimates of the global losses attributable to stem rust in a counterfactual 1961-2009 world where irrigated and high-input wheat areas were sown to rust susceptible wheat varieties[2], and d) the method we used to calculate the economically justifiable amount to spend on stem rust research given the counterfactual losses that were estimated to occur since 1961 if the world's wheat crop was susceptible to this pathogen.

1. Potential Spatial Distribution of Wheat Stem Rust

The potential distribution of wheat stem rust (*Puccinia graminis*) was estimated using the CLIMEX Compare Locations model (Sutherst 1985; Sutherst et al. 2007), run with the CliMond 1975H climatology (Kriticos et al. 2012). For the 'Compare Locations' analyses, CLIMEX requires five climatic variables in the form of long-term monthly means of daily minimum temperature, daily maximum temperature, relative humidity at 0900 hours and 1500 hours, and monthly total precipitation. The climate data used in the CLIMEX modelling was the CliMond 10' global historical dataset centred on 1975 (CM10_1975H_V1_WO (Kriticos et al. 2012). This dataset includes monthly average data for minimum and maximum air temperature, precipitation, and relative humidity recorded at 09:00 and 15:00. While the climate averages in the CM10_1975H_V1_WO dataset are centred on 1975, they include data from 1951-2000 for precipitation, and 1961-1990 for temperature (Hijmans et al. 2005) and humidity (New et al. 2002). To calculate a soil moisture balance and growth and stress indices, CLIMEX interpolates the monthly climate values down to weekly and daily values.

[1] Wheat is infected by three species of rusts: stem (black) rust (*Puccinia graminis* Pers. f. sp. *tritici*), leaf (brown) rust (*P triticina*), and stripe (yellow) rust (*P. striiformis* Westend. f. sp. *tritici*), but stem rust is deemed by some to have historically been the most devastating wheat disease worldwide (see, for example, Saari and Prescott 1985 and Singh et al. 2011). These rusts have complex asexual and sexual stages during their life cycles that typically span two host plants: urediospores produce successive generations asexually about once a week over summer on their (wheat) host and less rapidly on off-season volunteer wheats; telial spores survive over winter before germinating to produce basidiospores that infect the alternate host and on which sexual recombination occurs. The popular names for these rusts reflect their appearance during the telial stage for black rust and the uredinial stage for brown and yellow rust. Stakman (1957, p. 259) noted "There are more than three thousand kinds of plant rusts that attack grains and grasses, trees and shrubs," while Agrios (2005) puts the tally at more than five thousand species.

[2] We assumed the measured yields in low-input areas that typically use wheat varieties susceptible to stem rust or lack economic access to fungicides that ameliorate losses from infected wheat already include rust-induced losses from pre-Ug99 strains. Thus, the analysis considers only high-input and irrigated areas.

CLIMEX is a niche model that has been used to model numerous plant pathogens (e.g., Lanoiselet et al. 2002; Yonow et al. 2004; Pinkard et al. 2010; Watt et al. 2011; Yonow et al. 2013). It performs more reliably than correlative species distribution models when projected into novel climate situations, such as those encountered during inter-continental projections (Webber et al. 2011). It allows the modeler to use different types of information besides geographical distribution data to inform the selection of parameter values (e.g., infection experiments and phenological observations), and to provide strong cross-validation. CLIMEX calculates several state variables to help in fitting the model, and to understand the factors limiting the range of the taxon being modeled. At each geographical location (in this case a ten arc minute, geo-referenced grid cell), an annual Growth Index (GI_A) quantifies the potential for population growth during the favorable season, and the Ecoclimatic Index (EI) quantifies the overall potential for year-round population persistence, discounting the GI_A with the stresses accumulated during the unfavorable seasons. In most niche modeling exercises, attention is focused on the EI patterns to distinguish the regions where the species can persist with associated economic and other consequences. For some pest species, long-distance dispersal dynamics are important in understanding pest impact, for example aphids such as *Rhopalosiphum padi* which vectors Barley Yellow Dwarf Virus (MacFadyen and Kriticos 2012). In these situations, an understanding of the long-distance metapopulation dynamics are critical to understanding pest risks, and this requires an understanding of those areas that are suitable for year-round persistence, and those that can support growth during part of the year only. *Puccina graminis* is such a species; $EI \geq 1$ indicates those areas where populations can persist year round, and areas with $EI = 0$ and $GI_A \geq 10$, where epidemics can only occur during favorable seasons.[3]

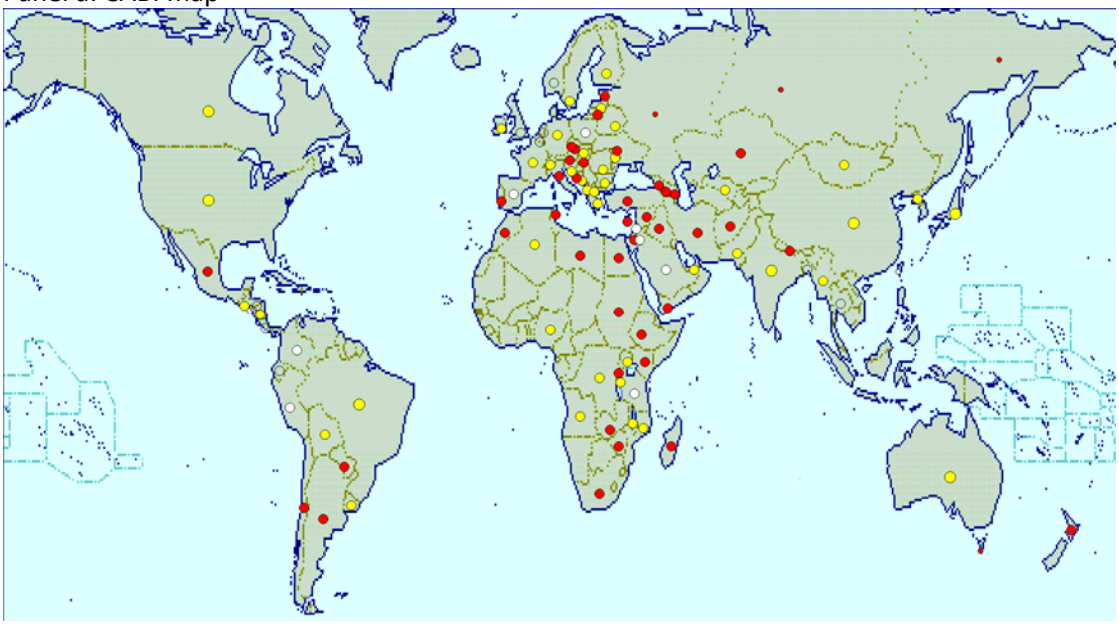
CLIMEX Compare Locations models usually rely on fitting model parameters to known suitable locations. The distribution of diseases is typically known only from very coarse scaled maps such as the country-level maps maintained by CABI (Figure 1, Panel a). For species niche modeling purposes, such maps are of limited utility due to the wide variety of climates that may occur within a single country. Such maps need to be linked to modeling exercises using fuzzy logic (at least one point in a known suitable country should be modeled as being climatically suitable, and countries from which the species is unknown are generally uninformative). The known distribution of wheat stem rust epidemics was assembled from various sources and compiled into a map (see Chai et al. 2013). In addition, experts were consulted to refine the map using an online spatial surveying tool (Beddow et al. 2010). An annual time-series of state-level occurrence data were available for the United States back to 1910. These spatial

[3] The thresholds were set to allow for at least one sexual cycle of *P. graminis*. Stakman (1957) noted that the pathogen multiplies rapidly, producing 50,000 to 450,000 urediniospores on a single plant within 7-10 days of infection.

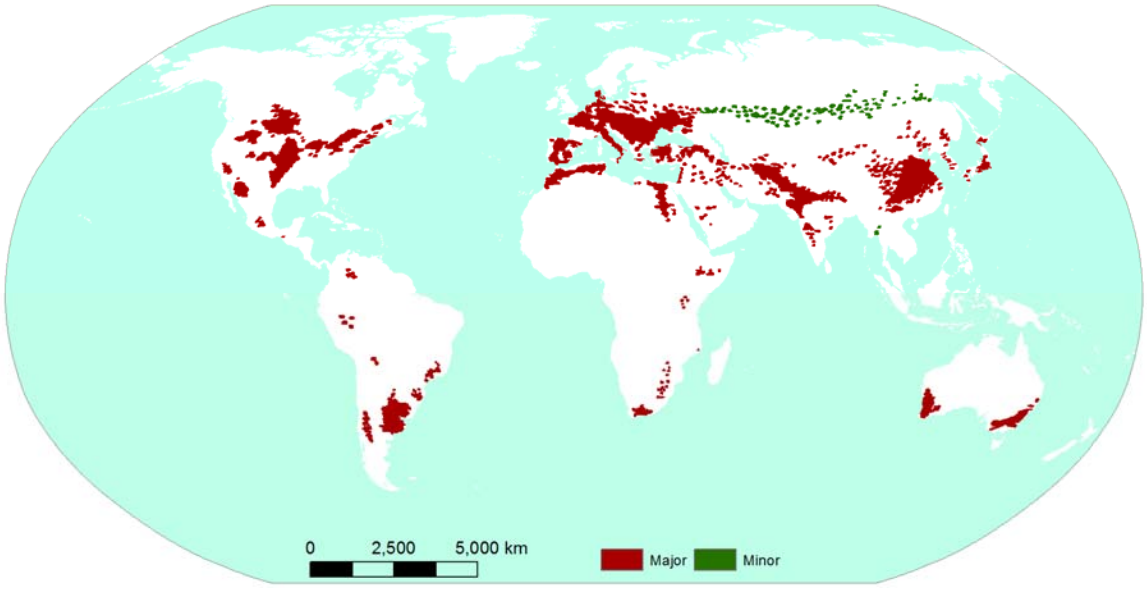
distribution data were supplemented by expert opinion, which was used to verify and validate the modeled and mapped results.

Figure 1: Existing Occurrence Maps for *Puccina graminis*

Panel a: CABI map



Panel b: Roelfs et al. map



Source: Panel a taken from CABI (2006). Panel b developed based on Roelfs, Singh and Saari (2002, p. 15).

Note: In Panel a, yellow dots indicate "present and no further details," red dots indicate "widespread, " and white dots indicate "present, localized" or "occasional or few reports."

Under a rainfed agricultural scenario, the model stresses were fitted to the known distribution where the *P. graminis* population overwinters. Parameters were fitted so that all known suitable locations were modeled as occurring in climatically suitable grid cells (Table 1; Figure 2). Cold stress parameters were adjusted to limit the northern extent of the persistent range in Texas, Louisiana and Alabama. The Dry Stress threshold was set to 0.1, to accord with the permanent wilting point. The Dry Stress accumulation rate was adjusted to limit the western extent of *P. graminis* in xeric locations such as New Mexico, though the Soil Moisture Growth Index was also limiting in this environment. Heat Stress was set to start at 34 °C, just above the maximum temperature for development, and the stress accumulation rate was fitted to the known distribution in the south-west of the United States. In the South-East of the United States, Hot Wet Stress parameters were adjusted to limit the potential range of *P. graminis* to known suitable locations. The underlying ecophysiological basis for this climatic stress is unknown, though it is likely to be associated with some form of biotic stress affecting the wheat host.

The Temperature Index parameters were set in consideration of the observed minimum temperature for infection (4 °C) reported by Tollenar (1985), and the upper maximum temperature reported by Kramer and Eversmeyer (1992), after accounting for the averaging effect involved in creating climatic data. The optimum temperatures were set in consideration of Rowell (1984), Tolenaar (1985) and Kramer and Eversmeyer (1992).[4] The Soil Moisture Index parameters were fitted to match observed patterns of abundance of *P. graminis* epidemics in the United States. The lower soil moisture limit (SM0) was set to 0.3, limiting the suitability in the mid-west of the United States under natural rainfall wheat cultivation. This value is well above permanent wilting point, which occurs when the soil moisture index is approximately 0.1. The relatively high value for SM0 suggests that the lower limit for population growth may be associated with conditions for spore germination and infection or that *P. graminis* may require stem turgor for growth. The lower and upper soil moisture index values for optimal growth spanned a relatively wide range around field capacity, reflecting the observed patterns of abundance in the United States. The upper soil moisture limit for growth was set to 2.5, reflecting conditions that are too wet for wheat persistence.

Two separate CLIMEX scenarios were run, one for irrigated wheat and another for rainfed wheat. The irrigated scenario assumed that 2.5 mm per day of top-up irrigation was applied during the winter months to identify regions where irrigated wheat would be climatically

[4] Rowell (1984) found that temperatures of 18 to 21 °C were ideal for sporulation and 15 to 24 °C were optimal for spore germination and appressorium formation. Tollenaar (1985) reported maximal germination of *P. graminis* spores at 20 °C, while Kramer and Eversmeyer (1992) reported maximal rates of germ tube elongation for *P. graminis* at 22 and 25 °C.

suitable for *P. graminis* epidemics (i.e., $GI_A \geq 10$). The results of each scenario were calculated globally for all land areas on the ten arc minute climatic grid.

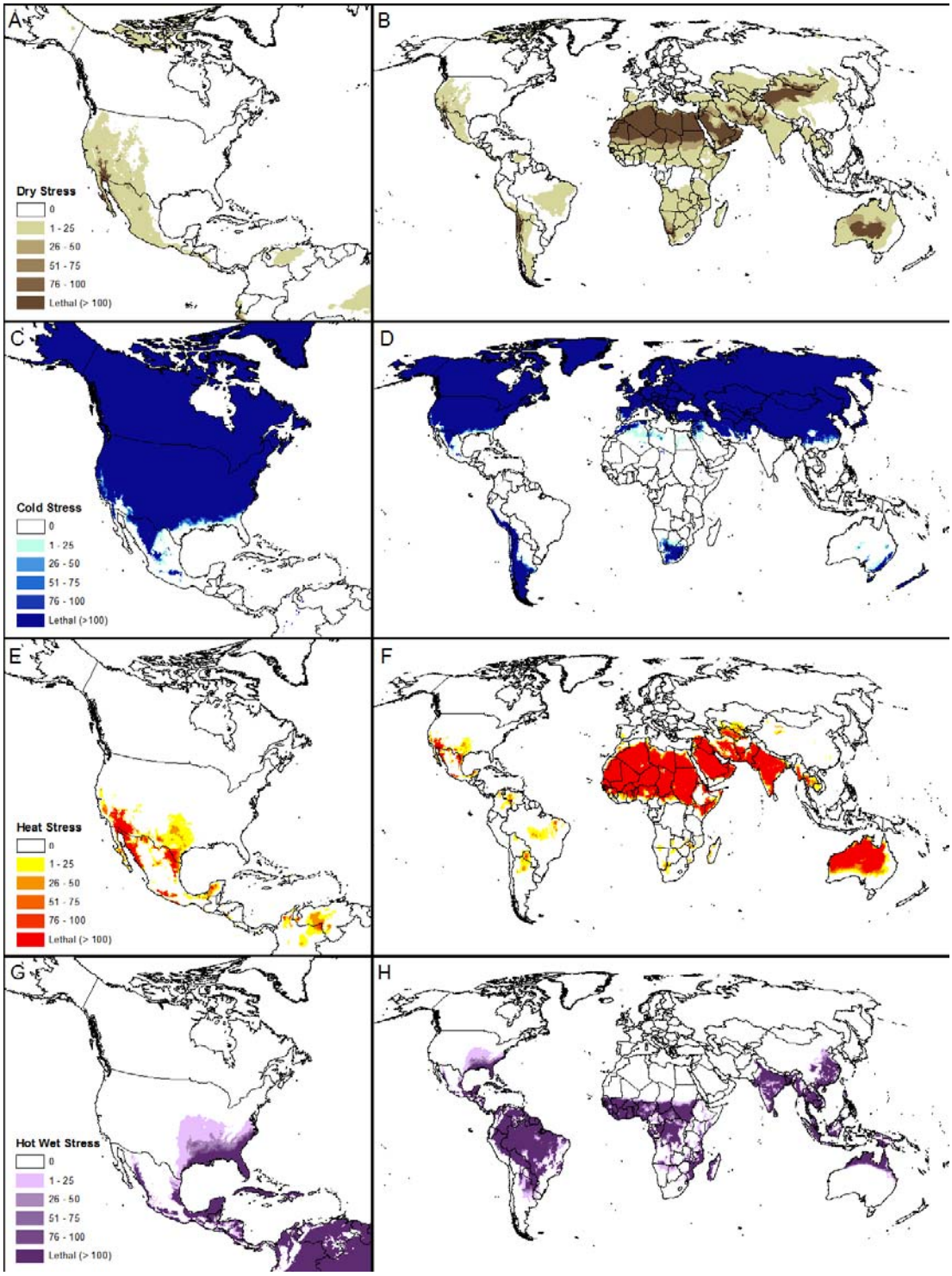
Table 1: CLIMEX parameter values used for modeling the distribution of *Puccina graminis*

	Parameter			
	Mnemonic	**Details**	**Value**	**Units[a]**
Temperature	DV0	lower threshold	4	°C
	DV1	lower optimum temperature	17	°C
	DV2	upper optimum temperature	25	°C
	DV3	upper threshold	32	°C
Moisture	SM0	lower soil moisture threshold	0.3	
	SM1	lower optimum soil moisture	0.6	
	SM2	upper optimum soil moisture	1.5	
	SM3	upper soil moisture threshold	2.5	
Cold stress	TTCS	damaging low threshold temperature	4	°C Days
	THCS	stress accumulation rate	-0.01	$Week^{-1}$
Heat Stress	TTHS	damaging high threshold temperature	34	°C
	THHS	stress accumulation rate	0.005	$Week^{-1}$
Dry stress	SMDS	soil moisture dry stress threshold	0.1	
	HDS	tress accumulation rate	-0.005	$Week^{-1}$
Hot Wet stress	TTHW	Hot wet temperature threshold	30	°C
	MTHW	soil moisture wet stress threshold	0.35	
	PHW	stress accumulation rate	0.03	$Week^{-1}$

Note: The role and meaning of parameters are described in Sutherst et al. (2007). See also Chai et al. (2013) for additional details.

[a] Where values are absent, units are a dimensionless index of available soil moisture, scaled from 0 (oven dry) to 1 (field capacity).

Figure 2: Stress Indices of *Puccina graminis*



Source: Created by authors.

Note: Each of the rows shows a different type of stress (dry, cold, heat and hot-wet) that limits the distribution of *Puccinia graminis*. The left column shows the average North American distribution of the stress while the right column shows its global distribution.

2. Wheat Crop Geography[5]

The HarvestChoice Spatial Allocation Model (SpAM) estimates a relatively high-resolution (five arc minute) global raster of production for a number of crops (You, Wood and Wood-Sichra 2006). For each crop and pixel, the model provides spatially disaggregated estimates of yield, output and area in each of three production systems (irrigated, high-input rainfed and low-input rainfed). The SpAM model output data (Version 3, Release 2) used in the present study are benchmarked on year 2000 production, and were spatially aggregated to 10 arc minute resolution to match the CLIMEX output described above[6]. To estimate the overall risk areas from *P. graminis,* the CLIMEX state variables were spatially intersected with the wheat crop geography information from SpAM, to create an integrated risk map, drawing on the appropriate CLIMEX scenario for the natural rainfall and irrigated wheat growing areas.

The world was divided into fifteen epidemiological zones (Figure 3) drawing on the ideas broached in Saari and Prescott (1985). The zones were constructed such that, in any given year, an *epidemic* in a given zone is likely to occur independently of epidemics in other zones due to the independence of climatological patterns between the zones. This does not imply that rust spores cannot spread between zones, but that spatio-temporal epidemic dynamics in one zone are unaffected by the dynamics in any other zone[7]. The main goal was to calculate the area harvested and the amount of production in each zone that is climatically suitable for stem rust infection and persistence. The CLIMEX results were transformed into a binary indicator of stem rust climatic suitability and potential for persistence in each 10 arc minute grid cell (see previous section). The portion of each zone, z, that is susceptible to stem rust was calculated as:

$$\beta_z = \left[\mathbf{a}_z'(\mathbf{GI}_{z,i} \circ \mathbf{I}_z) - \mathbf{a}_z'(\mathbf{GI}_{z,n} \circ (\mathbf{I}_z - \mathbf{1})) \right] / \mathbf{a}_z'\mathbf{1}$$

where

[5] This section draws from Beddow et al. (2010).

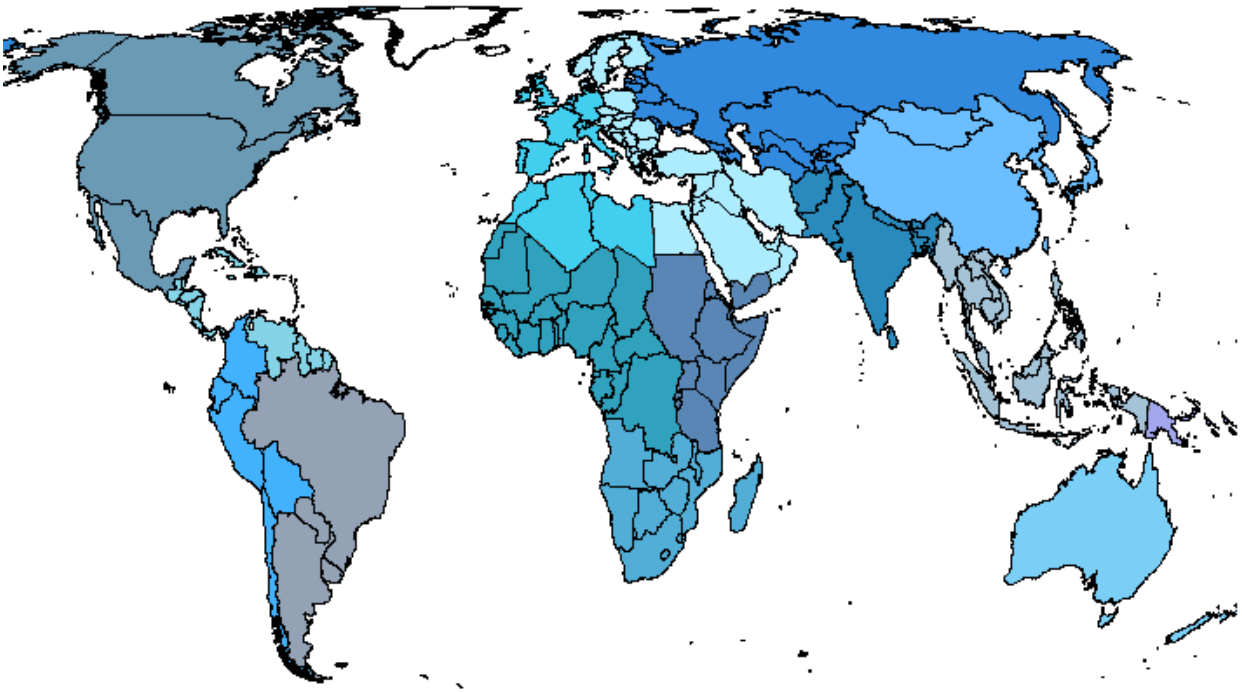
[6] The HarvestChoice crop geography data can be obtained from www.mapspam.info

[7] It seems plausible that at the local level the occurrence of stem rust in one location, for example (part of) a wheat field, will be highly correlated with its occurrence in another, nearby location (say another part of the same field). However, given the complex spatio-temporal stochastic processes affecting the development, dispersal, deposition and propagation of rust spores, there are good reasons to believe that as the scale of analysis increases, the occurrence of stem rust in one location is less likely linked with the occurrence of stem rust in another (even adjacent) location. For example, Appendix Figure 1 shows a selection of annual observations of reported stem rust occurrence state-by-state for the United States, and there is no apparent concordance of the occurrence of stem rust in one state versus another (even adjacent) state. This gives credence to our assumption of independence in stem rust occurrence among the zones.

- $\mathbf{a}_z$ is a vector of the total irrigated and high-input rainfed wheat production area in zone z's cells (as estimated by the SpAM model)
- $\mathbf{GI}_{z,i}$ is a vector of binary susceptibility indicators under the irrigated scenario for the same region (1 if suitable, 0 otherwise))
- $\mathbf{GI}_{z,n}$ is a vector of binary susceptibility indicators under the non-irrigated scenario (1 if suitable, 0 otherwise)
- $\mathbf{I}_z$ is a vector of binary indicators set to one if any wheat area in the corresponding cell was irrigated and zero otherwise
- $\mathbf{1}$ is a vector of ones of the appropriate length, and "$\circ$" is the operator indicating an element-by-element vector product (i.e., a Hadamard product).

The total amount of susceptible production and the total area and production under which stem rust can persist were derived similarly.

Figure 3: Epidemiological Zones Worldwide



Note: Appendix Table 2 reports a tabulation of countries grouped into epidemiological zones.

3. Probabilistic Estimates of Counterfactual Production Losses

Probabilistic estimates of counterfactual production losses were constructed separately for each epidemiological zone. Country-specific wheat production and area of production estimates taken from FAOSTAT were aggregated across these 15 zones for 1961 to 2009. This information was used to calculate the average yield per hectare by dividing the quantity of grain production by area of production. These average yield estimates were then adjusted to reflect disease-free yields in areas suitable for stem rust where stem rust resistant varieties have been used to control the *P. graminis* epidemics:

$$Y_{tz}^{p}\left(l_{tz}^{Post}\right)=\frac{Y_{tz}^{a}}{\left(1-\alpha_{tz}\right)+\alpha_{tz}\left(\beta_{z}\left(1-l_{tz}^{Post}\right)+\left(1-\beta_{z}\right)\right)}$$

where $Y_{tz}^{p}\left(l_{tz}^{Post}\right)$ is the average yield that would have been observed in year *t* and zone *z* if rust resistant wheat varieties eliminated all losses; α_{tz} is the proportion of zone *z* planted with rust resistant wheat varieties in year *t*; β_z is the proportion of the wheat producing area in zone *z* that is suitable for stem rust (see the previous section); and l_{tz}^{Post} is the loss in zone *z* in year *t*. This calculation assumes that stem rust losses from 1961 to 2009 were only avoided in areas where rust resistant wheat was planted and it was suitable for stem rust. With the exception of NorthEast Asia (mainly China), data for α_{tz} were taken from HarvestChoice's wheat crop geography map (Version 3, Release 2) based on high management and irrigated wheat production areas[8]. For NorthEast Asia, supplementary data for China were available from Dalrymple (1986), Byerlee and Moya (1993), Fan and Pardey (1992), Byerlee (1994), CIMMYT (1996), and Pingali, (1999) to better take into account the slower rate of adoption of rust resistant wheat varieties observed in China. The strategy used for obtaining l_{tz}^{Post} is discussed below. With $Y_{tz}^{p}\left(l_{tz}^{Post}\right)$ in hand, counterfactual average yields were calculated as

$$Y_{tz}^{cf}\left(l_{tz}^{Post},l_{tz}^{Pre}\right)=\left(\left(1-\alpha_{tz}\right)+\alpha_{tz}\left(\beta_{z}\left(1-l_{tz}^{Pre}\right)+\left(1-\beta_{z}\right)\right)\right)Y_{tz}^{p}\left(l_{tz}^{Post}\right)$$

where l_{tz}^{Pre} is the proportional yield loss that would have been observed in these regions without stem rust resistant wheat varieties.

These counterfactual yields can be differenced with actual yields and multiplied by area of production to measure counterfactual losses:

$$L_{tz}^{cf}\left(l_{tz}^{Post},l_{tz}^{Pre}\right)=\left(Y_{tz}^{a}-Y_{tz}^{cf}\left(l_{tz}^{Post},l_{tz}^{Pre}\right)\right)A_{tz}$$

[8] The SpAM data indicate that 55 percent of the Australasian wheat area is high-input. This was deemed too low and so the high-input area share for Australasia was set at 86 percent, equivalent to the respective area share in Southwest Asia.

where A_{tz} is the area of wheat production in zone *z* and year *t*. These losses were aggregated across regions:

$$L_t^{cf}\left(l_t^{Post}, l_t^{Pre}\right) = \sum_{z=1}^{15} L_{tz}^{cf}\left(l_{tz}^{Post}, l_{tz}^{Pre}\right)$$

where l_t^{Post} and l_t^{Pre} are vectors of losses across regions. Summing again over time yields cumulative counter factual losses:

$$L^{cf}\left(l^{Post}, l^{Pre}\right) = \sum_{t=1961}^{2009} L_t^{cf}\left(l_{tz}^{Post}, l_{tz}^{Pre}\right)$$

where l^{Post} and l^{Pre} are now vectors of losses across regions and years.

The challenge faced in calculating $L_t^{cf}\left(l_t^{Post}, l_t^{Pre}\right)$ and $L^{cf}\left(l^{Post}, l^{Pre}\right)$ is that we only have loss estimates with and without stem rust resistant wheat varieties for the United States[9]. Therefore, we used these reported losses to estimate the distribution of stem rust losses with and without resistant wheat varieties. These estimated distributions were used with Monte Carlo methods to construct the distribution of losses that could have occurred if the distribution of losses in each of the epidemiological zones was the same across zones, but observed losses could differ across zones in any given year. Effectively, these assumptions imply that stem rust losses are independently and identically distributed across epidemiological zones and time.

The annual proportion of wheat lost to stem rust in the United States from 1918 to 2009 was acquired from USDA (2011). These data were divided into two time periods, 1918 to 1960 and 1961 to 2009. From 1918 to 1960, U.S. wheat losses averaged 2.5 percent per year; losses dropped to less than an eighth of that rate (0.3 percent per year) thereafter. Some annual loss was recorded for every year from 1918 to 1960 resulting in a complete time series that could be

[9] In the United States, we estimate that 71.2 percent of the country's wheat area is suitable for stem rust but in only 0.9 percent of the area will it persist inter-seasonally. Elsewhere in the world the spatial concordance between the sink (i.e., susceptible or climatically suitable area) and source (i.e., the area where stem rusts persists) is much higher. For example, in sub-Saharan Africa, 82.5 percent of the wheat area is susceptible and in 61.0 percent of the area will the disease persist. In India the corresponding shares are 65.3 susceptible, 2.9 percent persistent. The United States is the only country known to us with the detailed, long-run yield loss data available in a form that facilitates implementation of the probabilistic loss method we have developed here. Given the physical separation between source and sink areas in the United States, using U.S. loss data to estimate the probable crop losses attributable to stem rust worldwide is likely to give us a conservatively low estimate of those global losses. To explore the prospective upper bounds of these global losses, Pardey et al. (2013) took the worst case scenario that occurred in the United States during the past century--specifically 1935 when the United States lost 24 percent of its wheat crop to stem rust--and assumed a pandemic outbreak in 2010 wherein the entire world's susceptible wheat area experienced a reduction in wheat yields equivalent to the outlier 1935 event in the United States. They found that even this worst case scenario, applied carefully to the world's susceptible areas, produced global loss estimates that were much more muted than those hitherto reported.

tested for non-stationarity (e.g., augmented Dickey Fuller test) and autocorrelation (e.g., Bartlett test for white noise), which were both rejected. Data for the second period (1961 to 2009), also from USDA (2011), are more challenging since no losses were recorded in almost one out of three years. Furthermore, when losses were recorded, they were typically quite small. Therefore, we chose to model losses as a conditional beta distribution with no autocorrelation:

$$f(l)=\begin{cases} p & \text{for } l=0 \\ (1-p)\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} l^{a-1}(1-l)^{b-1} & \text{otherwise} \end{cases}$$

where $1 > l \geq 0$ is the reported proportional loss, p is the probability that no loss was reported, $\Gamma(\cdot)$ is the gamma function, and $a > 0$ and $b > 0$ are the parameters of the beta distribution conditional on a positive loss. The parameters p, a, and b were estimated using maximum likelihood. This was accomplished by programming the likelihood function in STATA (version 11.2) and then maximizing it using STATA's ml max command[10]. The resulting parameter estimates are reported in Table 2, while Figure 4 shows the fit of these estimates.

Table 2: Conditional Beta Parameter Estimates (standard errors) for Reported U.S. Stem Rust Losses by Period.

Parameter	1918 to 1960	1961 to 2009
p	0.0[a]	0.327***
		(0.067)
a	0.532***	0.305***
	(0.095)	(0.059)
b	20.33***	67.62***
	(5.50)	(24.91)
Maximized Log-likelihood	121.78	141.43
Observations	43	49

[a] Standard error could not be estimated. *** Significant at 0.01.

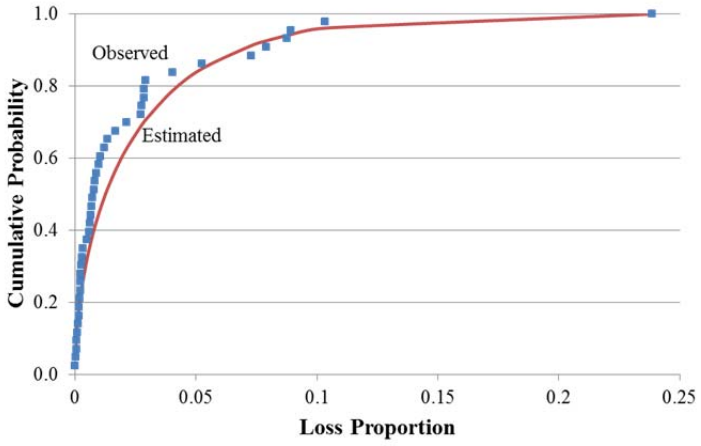
[10] See Appendix 2 for the source code used to estimate the parameters of the beta distributions.

Monte Carlo Distributions

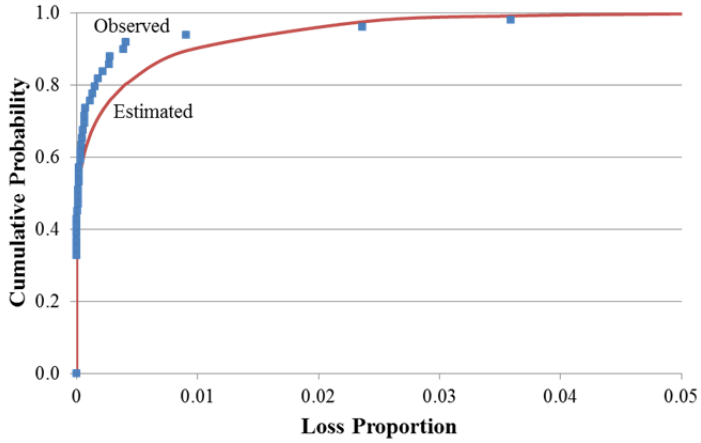
The distributions for $L_t^{cf}\left(l_t^{Post}, l_t^{Pre}\right)$ and $L^{cf}\left(l^{Post}, l^{\text{Pre}}\right)$ were constructed by randomly drawing two uniform variates for each zone and time period, and for 50,000 replicates: u_{tz}^r and v_{tz}^r where *r* represents the replicate. The first random variate was used to determine whether or not the zone experienced a loss. If $u_{tz}^r \geq p$, a yield loss in zone *z* and period *t* was assumed to have occurred with rust resistant varieties. The size of the loss if one occurred was calculated as $l = beta^{-1}(v_{tz}^r, a, b)$. For l_{tz}^{Post}, the parameter estimates for 1961 to 2009 were used. For l_{tz}^{Pre}, the parameter estimates for 1918 to 1960 were used. This Monte Carlo simulation was implemented using Microsoft Excel 2010 with the Palisade @Risk (version 5.0.0 Professional Edition) add-in. Figure 5 shows the cumulative distribution of counterfactual losses from 1961 to 2009 based on this method.

Figure 4: Distribution of Observed and Fitted Stem Rust Losses

Panel a: 1918 to 1960

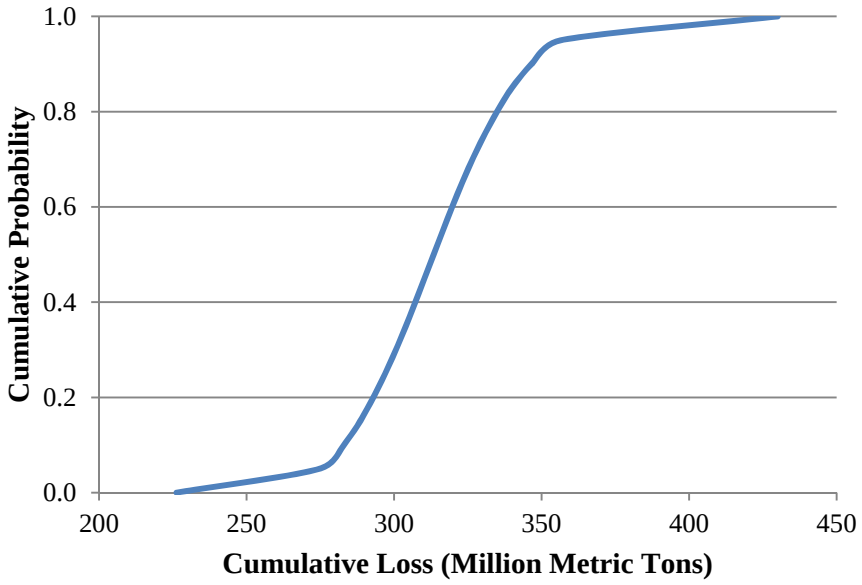


Panel b: 1961 to 2009



Source: Observed data from USDA (2011).

Figure 5: Cumulative Distribution of Counterfactual Stem Rust Losses from 1961 to 2009



Source: Estimated by authors.

4. Economically Justified Wheat R&D

The economically justified annual investments in wheat R&D were calculated based on the annual distribution of counterfactual wheat losses. We defined this economically justified investment in two different ways:

$$V\left(r^{i}\right)=\max\left\{v:\Pr\left(\sum_{t=1961}^{2009}p_{t}L_{t}^{cf}\left(l_{tz}^{Post},l_{tz}^{Pre}\right)\left(1+r^{i}\right)^{2009-t}\geq\sum_{t=1951}^{2009}v\left(1+r^{i}\right)^{2009-t}\right)>0.95\right\}\text{ and}$$

$$V\left(r^{r},r^{m}\right)=\max\left\{v:\Pr\left(\sum_{t=1961}^{2009}p_{t}L_{t}^{cf}\left(l_{tz}^{Post},l_{tz}^{Pre}\right)\left(1+r^{r}\right)^{2009-t}\geq\sum_{t=1951}^{2009}v\left(1+r^{m}\right)^{2009-t}\right)>0.95\right\}$$

where p_t is the price of wheat in year t, r^i is the internal rate of return, r^r is the reinvestment rate, and r^m is the modified internal rate of return. The first equation yields the maximum value that could have been invested annually from 1951 to 2009 while still achieving an internal rate of return of at least r^i with a probability of 0.95, and assuming that the counterfactual distribution of losses was averted. The second equation yields the maximum value that could have been invested annually from 1951 to 2009 while still achieving a modified internal rate of return of at least r^m with a probability of 0.95 given a reinvestment rate of r^r and assuming the counterfactual distribution of losses was averted. These calculations assume that the initial R&D investment takes ten years before beginning to yield a return. They were also performed with @RISK and its goal seek function using U.S. wheat prices (USDA-ERS) deflated to 2010

values using the consumer price index (BLS 2011), an internal rate of return of 20 percent per year, a modified internal rate of return of 10 percent per year, a cost of capital discount rate equal to the modified internal rate of return, and a reinvestment rate of 3 percent per year. An internal rate of return of 20 percent per year is consistent with the average internal rate of return found in the recent study by Alston et al. (2010). An average modified internal rate of return of about 10 percent per year with a reinvestment rate of 3 percent per year was also found in Alston et al. (2011) and Rao et al. (2012)[11]. The Rao et al. estimate is derived from 300 evaluations that encompassed a portfolio of agricultural research from around the world. Thus, it provides a suitable economic benchmark for the opportunity cost of research directed to wheat stem rust vis-à-vis other global agricultural R&D[12].

Using the MIRR approach, we estimate that a sustained investment of $51.1 million per year (2010 prices) in stem rust research could be economically justified. Dividing this investment by the total hectares of wheat produced in 2009 and the total value of this production yields an investment of $0.23 per year per hectare or $29.44 per hundred thousand dollars of wheat production. By comparison, benchmarking these calculations on a conventional IRR of 20 percent per year suggests that annual investments on stem rust research of $159.2 million are economically justifiable ($0.71 per year per hectare or $97.72 per hundred thousand dollars of annual wheat production).

[11] Rao et al. (2012) report that the mean internal rate of return from a sample of 300 observations obtained from studies of the impacts of agricultural R&D undertaken around the world since the 1950s averaged 52 percent per year, which implied an average MIRR of around 20 percent per year. As Alston et al. (2010) described, there are good reasons to believe many of these IRR and MIRR estimates are upwardly biased due to model misspecification issues, and indications that the specification choices (especially allowing for longer R&D lags) made in more recent studies are likely to reduce the bias. The data underlying Rao et al. (2012) are consistent with this notion: the IRR for studies done in the 1950s and 1960s averaged 78 percent per year, versus 36 percent per year for studies done since 2000. For these reasons we chose an IRR of 20 percent per year and a MIRR of 10 percent per year to benchmark our estimates of the economically justifiable amount of spending on stem rust research worldwide.

[12] We consulted knowledgeable individuals and undertook a meta-review of the wheat research literature to form a judgment about the global public investments in wheat stem rust research against which to compare the R&D investments we deem are economically justified given the crop loss estimates developed by way of this study.

References

Agrios, G. (2005). *Plant Pathology* 5th Edition, Boston: Elsevier Academic Press.

Alston, J.M., M.A. Andersen, J.S. James, and P.G. Pardey (2010). *Persistence Pays: U.S. Agricultural Productivity Growth and the Benefits from Public R&D Spending.* New York: Springer.

Alston, J.M., M.A. Anderson, J.S. James and P.G. Pardey (2011). The economic returns to U.S. public agricultural research. *American Journal of Agricultural Economics* 93(5):1257-1277.

Beddow, J.M., D.J. Kriticos, P.G. Pardey and R.W. Sutherst (2010). "Potential Global Crop Pest Distributions Using CLIMEX: *HarvestChoice* Applications." HarvestChoice Working Paper. St Paul: University of Minnesota.

Bureau of Labor Statistics (BLS) (2011). Consumer Price Index History Table. Washington, D.C.: Bureau of Labor Statistics. Accessed November 22, 2011 from http://www.bls.gov/cpi/#tables.

Byerlee, D. and P. Moya (1993). *Impacts of International Wheat Breeding Research in Developing World, 1966-1990*. Mexico, D.F., Centro Internacional de Mejoramiento de Maiz y Trigo.

Byerlee, D. (1994). *Modern Varieties, Productivity, and Sustainability: Recent Experience and Emerging Challenges*. Mexico: Centro Internacional de Mejoramiento de Maiz y Trigo.

CABI. *CABI Crop Protection Compendium*. Data Disk. Wallingford, UK: CAB International, 2006.

Chai, Y., J. Beddow, R.W. Sutherst, E. Duvillier and D. Kriticos (2013). *Wheat Stem Rust Pest Geography*. HarvestChoice. St Paul: University of Minnesota, in press.

CIMMYT (Centro Internacional de Mejoramiento de Maiz y Trigo) (1996). *CIMMYT 1995/96 World Wheat Facts and Trends: Understanding Global Trends in the Use of Wheat Diversity and International Flows of Wheat Genetic Resources*. Mexico: Centro Internacional de Mejoramiento de Maiz y Trigo..

Dalrymple, D.G. (1986*). Development and Spread of High-Yielding Wheat Varieties in Developing Countries*. Washington, D.C.: U.S. Agency for International Development.

Fan, S. and P.G. Pardey (1992). *Agricultural Research in China*. The Hague: International Service for National Agricultural Research.

Fisher, M.C., D.A. Henk, C.J. Briggs, J.S. Brownstein, L.C. Madoff, S.L. McCraw and S.J. Gurr (2012). Emerging fungal threats to animal, plant and ecosystem health. *Nature* 484:186-194.

Hijmans R.J., S.E. Cameron, J.L. Parra, P.G. Jones, and A. Jarvis (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965-1978.

Hurley, T.H. (2010). A review of agricultural production risk in the developing world. HarvestChoice Working Paper. St Paul: University of Minnesota.

Kramer C.L. and M.G. Eversmeyer (1992). Effect of temperature on germination and germ-tube development of *Puccinia recondita* and *P. graminis* urediniospores. *Mycological Research* 96:689-693.

Kriticos D.J., B.L. Webber, A. Leriche, N. Ota, J. Bathols, I. Macadam and J.K. Scott (2012). CliMond: global high resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods in Ecology and Evolution* 3:53-64.

Lanoiselet, V., E.J. Cother, and G.J. Ash (2002). Climex and Dymex simulations of the potential occurrence of rice blast disease in south-eastern Australia. *Australasian Plant Pathology*, 31:1-7.

Macfadyen, S. and D.J. Kriticos (2012). Modelling the geographical range of a species with a variable life-history. *Public Library of Science One* 7(7).

New M., D. Lister, M. Hulme, and I. Makin (2002). A high-resolution data set of surface climate over global land areas. *Climate Research* 21: 1-25.

Pardey, P.G., J.M. Beddow, D.J. Kriticos, T.M. Hurley, R.F. Park, E. Duveiller, R.W. Sutherst, J.J. Burdon, and D. Hodson (2013). Right-sizing stem-rust research. *Science* 340:147-148.

Peterson P.D., ed. (2001). *Stem Rust of Wheat: From Ancient Enemy to Modern Foe*. St. Paul: APS Press.

Pingali, P.L. (1999). *CIMMYT 1988-99 World Wheat Facts and Trends. Global Wheat Research in a Changing World: Challenges and Achievements*. Facts and Trends/Overview and Outlook No. 23726. Mexico: Centro Internacional de Mejoramiento de Maiz y Trigo.

Pinkard, E.A., D.J. Kriticos, T.J. Wardlaw and A.J. Carnegie (2010) Estimating the spatio-temporal risk of disease epidemics using a bioclimatic niche model. *Ecological Modelling*. 221 (23):2828-2838.

Rao, X., T.M. Hurley and P.G. Pardey (2012). Recalibrating the reported rates of return to food and agricultural R&D. Department of Applied Economics Staff Paper P12-8. St Paul: University of Minnesota.

Roelfs, A.P., R.P. Singh and E.E. Saari (1992). *Rust Diseases of Wheat: Concepts and Methods of Disease Management*. Mexico: Centro Internacional de Mejoramiento de Maiz y Trigo.

Rowell J.B. (1984). Controlled infection by *Puccinia graminis* f. sp. *tritici* under artificial conditions. In W.R. Bushnell and A.P.Roelfs, eds. *The Cereal Rusts: Origins, Specificity, Structure, and Physiology, Vol. I*. Orlando: Academic Press.

Saari E.E. and J.M. Prescott (1985). "World distribution in relation to economic losses." Chapter 9 in A.P. Roelfs and W.R. Bushnell, eds., *The Cereal Rusts: Diseases, Distribution, Epidemiology and Control, Vol. II.* Orlando: Academic Press.

Singh R.P., D. P. Hodson, J. Huerta-Espino, Y. Jin, S. Bhavani, P. Njau, S. Herrera-Foessel, P.K. Singh, S. Singh and V. Govindan (2011).The emergence of Ug99 races of stem rust fungus is a threat to world wheat production. *Annual Review of Phytopathology* 49: 465-481.

Stakman, E.C. (1957). Problems in preventing plant disease epidemics. *American Journal of Botany* 44: 259-267.

Sutherst R.W. and G.F. Maywald (1985). A computerised system for matching climates in ecology. *Agriculture, Ecosystems and Environment* 13:281-299.

Sutherst R.W., G.F. Maywald, and D.J. Kriticos (2007). *CLIMEX Version 3: User's Guide*. South Yarra: Hearne Scientific Software Pty Ltd (www.Hearne.com.au).

Tollenaar H. (1985). Uredospore germination and development of some cereal rusts from south-central Chile at constant temperatures. *Journal of Phytopathology* 114:118-125.

USDA, National Agricultural Statistics Service (2010). Crop Values 2009 Summary. Washington, D.C.: U.S. Department of Agriculture, NASS.

USDA, Agricultural Research Service (2011). Small grains losses due to rust Washington, D.C.: U.S. Department of Agriculture, ARS.(www.ars.usda.gov/Main/docs.htm?docid=10123). Accessed May 2011.

USDA, Economic Research Service (2012). Wheat Data: Yearbook Tables. Wheat: Average price received by farmers, United States. Table 18-- Wheat: Average price received by farmers, United States (dollars per bushel). Washington, D.C.: U.S. Department of Agriculture, ERS. (http://www.ers.usda.gov/data/wheat/YBtable18.asp).

Watt M.S., R.J. Ganley, D.J. Kriticos and L.K. Manning (2011). Dothistroma needle blight and pitch canker: the current and future potential distribution of two devastating diseases of *Pinus* species. *Forest Ecology and Management* 41(2):412-424.

Webber B.L., C.J. Yates, D.C. Le Maitre, J.K. Scott, D.J. Kriticos, N. Ota, A. McNeill, J.J. Le Roux and G.F.Midgley (2011). Modelling horses for novel climate courses: insights from projecting potential distributions of native and alien Australian acacias with correlative and mechanistic models. *Diversity and Distributions* 17(5):978-1000.

World Bank (2012). World Development Indicators. Downloaded March 22, 2012 from databank.worldbank.org.

Yonow T., D.J. Kriticos and R.W. Medd (2004). The potential geographic range of *Pyrenophora semeniperda*. *Phytopathology.* 94(8):805-812.

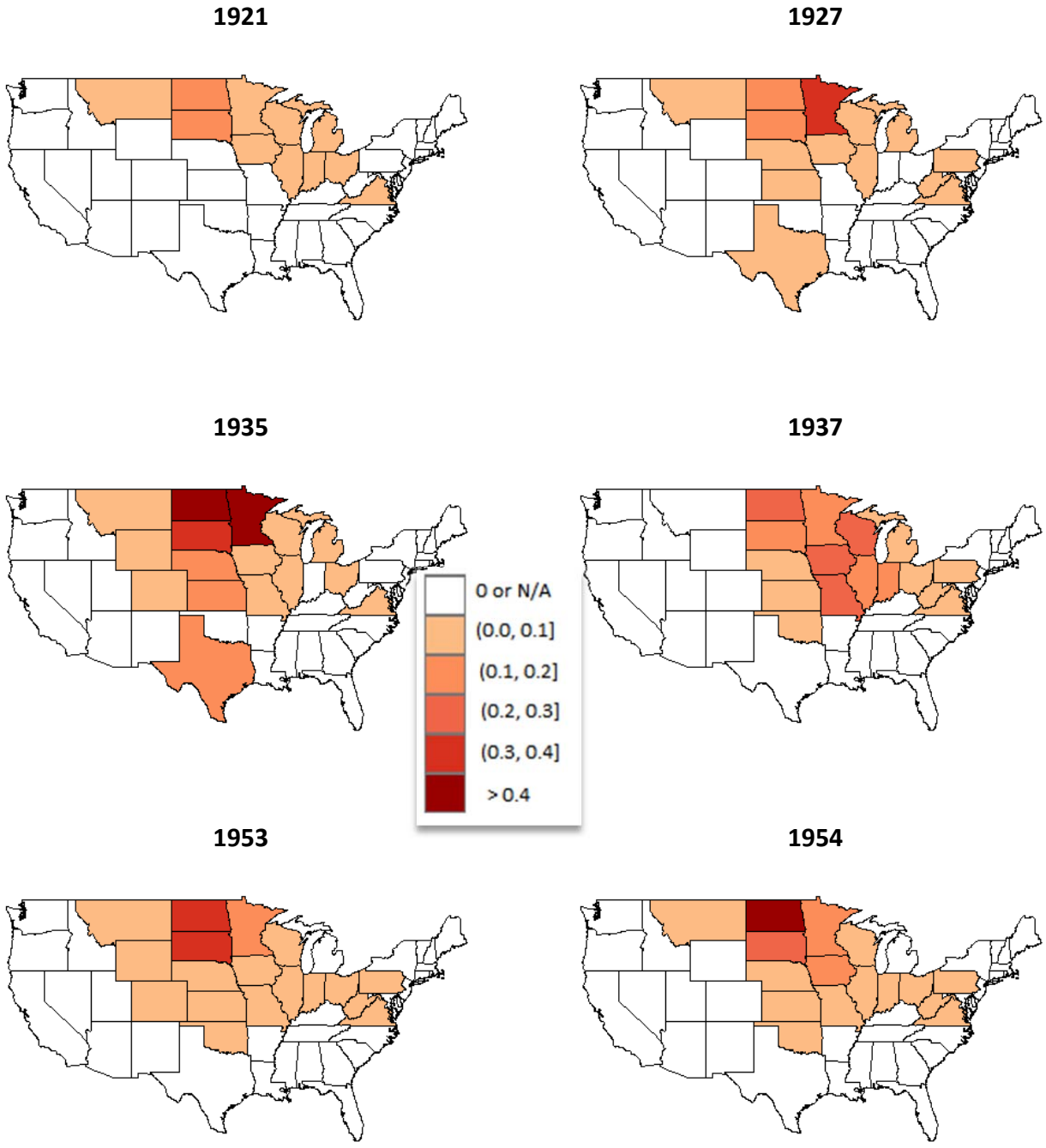
Yonow, T., V. Hattingh. and M. de Villiers (2013). CLIMEX modelling of the potential global distribution of the citrus black spot disease caused by *Guignardia citricarpa* and the risk posed to Europe. *Crop Protection* 44 (0):18-28.

You, L., S. Wood, and U. Wood-Sichra (2006). Generating global crop maps: from census to grid. Selected paper, IAAE (International Association of Agricultural Economists) Tri-annual Conference, Gold Coast, Australia.

Appendix: Undernourishment and Wheat Loss

USDA-ERS (2012) reports that 1 million kg of hard spring flour has a grain equivalent of 1.370 million kg. Thus, about 73 percent of wheat grain converts to flour. Beginning with an annual global wheat loss estimate of 6.2 MT, we therefore estimate that about 4.526 MT of wheat flour is lost, with a caloric value of 16.475 trillion kcal. World Bank (2012) reports that about 22 percent of the Sub-Saharan African population was undernourished in 2008, with a daily calorie deficit of 243 kcal per capita. Given a Sub-Saharan African population of about 854.227 million, this implies that the undernourished population is about 187.939 million, with a total annual calorie deficit of 16.681 trillion kcal. Thus, the caloric content of the 6.2 MMT annual wheat loss is almost sufficient to meet the Sub-Saharan African calorie deficit.

Appendix Figure 1: Reported State Losses (share) Due to Stem Rust in the U.S., Selected Years.



Source: Created by the authors using data from USDA, ARS (2011).

Appendix Table 1: Countries Grouped into Epidemiological Zones

Andean Region
Bolivia
Chile
Colombia
Ecuador
Peru

Australasia
Australia
New Zealand

Central America and Caribbean
Anguilla
Antigua and Barbuda
Aruba
Bahamas
Barbados
Belize
British Virgin Islands
Cayman Islands
Costa Rica
Cuba
Dominica
Dominican Republic
El Salvador
French Guiana
Grenada
Guadeloupe
Guatemala
Guyana
Haiti
Honduras
Jamaica
Martinique
Montserrat
Netherlands Antilles
Nicaragua
Panama
Puerto Rico
Saint Kitts and Nevis
Saint Lucia
Saint Vincent and the Grenadines
Suriname
Trinidad and Tobago
Turks and Caicos Islands
United States Virgin Islands
Venezuela

Eastern South America
Argentina
Brazil
Falkland Islands (Malvinas)
Paraguay
Uruguay

East Central Africa
Burundi
Djibouti
Eritrea
Ethiopia
Ethiopia PDR
Kenya
Rwanda
Seychelles
Somalia
Sudan
Uganda
United Republic of Tanzania
Yemen

Eurasia
Albania
Austria
Bahrain
Bosnia and Herzegovina
Bulgaria
Croatia
Cyprus
Czech Republic
Czechoslovakia
Denmark
Egypt
Finland
Greece
Hungary
Iran (Islamic Republic of)
Iraq
Israel
Jordan
Kuwait
Lebanon
Montenegro
Norway
Occupied Palestinian Territory
Oman
Poland
Qatar
Romania
Saudi Arabia
Serbia
Serbia and Montenegro
Slovakia
Slovenia
Sweden
Syrian Arab Republic
Fmr. Yugoslav Rep. of Macedonia
Turkey
United Arab Emirates
Yugoslav SFR

Former Soviet Union
Armenia
Azerbaijan
Belarus
Estonia
Georgia
Kazakhstan
Kyrgyzstan
Latvia
Lithuania
Republic of Moldova
Russian Federation
Tajikistan
Turkmenistan
Ukraine
USSR
Uzbekistan

North Africa and West Europe
Algeria
Andorra
Belgium
Belgium-Luxembourg
Channel Islands
Faroe Islands
France
Germany
Gibraltar
Holy See
Iceland
Ireland
Isle of Man
Italy
Libyan Arab Jamahiriya
Liechtenstein
Luxembourg
Malta
Monaco
Morocco
Netherlands
Portugal
San Marino
Spain
Switzerland
Tunisia
United Kingdom

North America
Bermuda
Canada
Greenland
Mexico
Saint Pierre and Miquelon
United States of America

Appendix Table 1 : Countries Grouped into Epidemiological Zones (continued)

Northeast Asia
China
North Korea
Japan
Mongolia
Republic of Korea

Pacific
American Samoa
Cook Islands
Fiji
French Polynesia
Guam
Kiribati
Marshall Islands
Micronesia (Federated States of)
Nauru
New Caledonia
Niue
Norfolk Island
Northern Mariana Islands
Pacific Islands Trust Territory
Palau
Papua New Guinea
Samoa
Solomon Islands
Tokelau
Tonga
Tuvalu
Vanuatu
Wallis and Futuna Islands

Southeast Asia
Brunei Darussalam
Cambodia
Indonesia
Lao People's Democratic Republic
Malaysia
Myanmar
Philippines
Singapore
Thailand
Timor-Leste
Viet Nam

Southern Africa
Angola
Botswana
Comoros
Lesotho
Madagascar
Malawi
Mauritius
Mozambique
Namibia
Réunion
Saint Helena
South Africa
Swaziland
Zambia
Zimbabwe

Southwestern Asia
Afghanistan
Bangladesh
Bhutan
India
Maldives
Nepal
Pakistan
Sri Lanka

West Central Africa
Benin
Burkina Faso
Cameroon
Cape Verde
Central African Republic
Chad
Congo
Côte d'Ivoire
D.R. Congo
Equatorial Guinea
Gabon
Gambia
Ghana
Guinea
Guinea-Bissau
Liberia
Mali
Mauritania
Niger
Nigeria
Sao Tome and Principe
Senegal
Sierra Leone
Togo
Western Sahara

Appendix 2. Stata source code to estimate parameters of the conditional beta distribution

```
program loglikelihood_beta

    version 10.1

    args lnf01  a b  probl


    tempvar  x PRB


    quietly{
           generate double `x'     = $ML_y1
            generate       `PRB'   = (`probl' ^ 2) /(1 + `probl' ^ 2)

            replace        `lnf01' = ln(1 - `PRB') + ln(betaden(exp(`a'), exp(`b'), `x')) if `x' >  0
            replace        `lnf01' = ln(`PRB')                                  if `x' <= 0
   }
end

cons   1 [probl]_cons = 0

* Estimate Conditional Beta for pre-1961 losses and calculate parameters values
ml model lf loglikelihood_beta (a:     proploss = ) (b:     )  (probl:     ) if year < 1961
ml max
nlcom (prb: ([probl]_cons ^ 2) /(1 + [probl]_cons ^ 2))
nlcom (a: exp([a]_cons))
nlcom (b: exp([b]_cons))

* Estimate Conditional Beta for post-1960 losses and calculate parameters values
ml model lf loglikelihood_beta (a:     proploss = ) (b:     )  (probl:     ) if year >= 1961
ml max
nlcom (prb: ([probl]_cons ^ 2) /(1 + [probl]_cons ^ 2))
nlcom (a: exp([a]_cons))
nlcom (b: exp([b]_cons))
```