WORKING PAPER

Potential Global Crop Pest Distributions Using CLIMEX: HarvestChoice Applications

Jason M. Beddow
Darren Kriticos
Philip G. Pardey
Robert W. Sutherst*

August 17, 2010

*Authorship is alphabetical.

Beddow is a graduate student and Pardey a professor, both at the Department of Applied Economics, University of Minnesota, St. Paul. Kriticos is a Principal Research Scientist with CSIRO in Australia. Sutherst is an Honorary Professor, The Ecology Centre, School of Integrative Biology, University of Queensland (formally Entomology Division, CSIRO).

This paper was prepared for the HarvestChoice project with support from the Bill and Melinda Gates Foundation.

HarvestChoice undertakes broad-ranging evaluation of technologies and strategies to inform policy and investment choices designed to raise the productivity of the agricultural systems most beneficial to the poor.

Executive Summary

HarvestChoice provides data, information and tools to support strategic investment decisions in agriculture. Most agricultural processes are strongly influenced by biotic stressors, such as insects and pathogens. However, available pest information has been lacking, for example global maps of where pests and diseases could occur have not been available. To address this deficiency, HarvestChoice has developed and implemented an approach to pest modeling aimed directly at supporting strategic decisions in agriculture. The HarvestChoice system includes methods, techniques and tools to both collect primary data on pest occurrence and to model potential global pest occurrence. In addition, a global team of recognized pest and climate experts has been assembled to aid in modeling, validation and peer-review.

The pest modeling system developed by HarvestChoice differs from the more typical monitoring and prediction systems that support short- to medium-term objectives, such as responding to pest events. Rather, we support strategic decision-making by mapping not where pests occur, but where they might potentially occur. This approach allows us to answer hypothetical questions that were previously unanswerable. For example, we can provide broad information on what portion of the world's crop area could be subject to a pest or what might happen if a new variety were planted in a previously un-cropped location.

The end results of our efforts are maps of potential pest occurrence. These maps are produced by modeling a species' response to factors such as temperature and moisture, while carefully keeping certain other factors exogenous (namely, those factors that might be decision variables for strategic planning). For example, the models are not conditional on the actual presence of a susceptible host, so our maps show the potential pest distribution assuming that susceptible hosts occur everywhere, avoiding confounding the host climate responses with those of the pest.

The models combine geo-spatial climate data and information on the seasonal phenology of each pest to investigate potential growth, stress and persistence of pests on a global scale. Growth is measured by a "Growth Index," higher values of which indicate a higher potential for that the pest's population to grow at a given location. Potential growth is restricted by sub-optimal temperatures or moisture and the pest's survival is limited by stresses, such as extremely cold or dry conditions. A final index, the "Ecoclimatic Index," is calculated by considering the extent to which growth is possible during the favorable season and survival is limited by the stresses during the non-growth season at each location. The result of this calculation, the Ecoclimatic Index is a measure of the potential ability of the pest to persist and develop high population sizes at a location: that is, whether it can survive the extremes of summer, winter, dry and wet seasons. This Working Paper describes the *HarvestChoice* approach to pest modeling and places that work in the context of other pest mapping efforts.

Potential Global Crop Pest and Disease Distributions
Using CLIMEX: *HarvestChoice* Applications

1. Introduction

The number of potential pest-crop combinations is large. CABI (2006) documented over 8,000 specific insects, diseases, weeds and other pests affecting the 2,191 beneficial plants in the CABI database, with an estimated 875 unique pests affecting just maize, wheat and rice.[1] Not all of these plant pests are of equal economic significance in terms of the frequency and spatial extent of their occurrence or the crop yield and quality losses (or costs of control) they incur. Absent information about the global extent and the biological-cum-economic consequences of these pests, decisions regarding the optimal investment in public and private efforts to avoid or mitigate these pest impacts are left largely to (informed) guesswork.

Conscious of the lack of information about the potential global extent of crop pests, *HarvestChoice* develops improved, evidenced-based approaches to setting strategic research priorities and choosing intervention options to address the productivity and production risk consequences of crop pests and diseases.[2] This paper reports progress to date, highlighting *HarvestChoice* efforts to develop a suite of geo-referenced maps of the potential global distribution of crop pests of suspected economic significance affecting the principal food and feed crops worldwide.

To simulate the potential for occurrence of the key pests and diseases we opted to use the CLIMEX model, working closely with leading plant pathologists and entomologists from around the world to calibrate and validate the maps. The general structure of the CLIMEX approach is described here, as well as the specific approach we took to generate a suite of maps of the key pathogens that affect cassava, wheat, maize and banana. We describe the data capture and processing methods we developed and used to calibrate and validate the pathogen maps, and give guidance as to their use and interpretation. This undertaking has produced a novel set of global, geo-referenced maps of potential crop pest and disease occurrences to inform strategic agricultural investment choices, crop research opportunities, and agricultural development decisions. We place this *HarvestChoice* effort in the context of prior efforts at a global level as well as contemporary efforts to monitor, track and forecast crop

---

[1] To avoid confusion, the term "pest" is used as a generic term to refer to any organism that generally has a negative influence on crop production, regardless of whether it is a pathogen, insect or other type of organism. When necessary, more specific terms are used.

[2] Pests not only affect yield levels (or crop productivity); they also affect the variance in crop yields and quality, and thus the riskiness of production. In a related *HarvestChoice* effort, Hurley (2010) provides a recent and thorough review and evaluation of the literature pertaining to production risk, highlighting the risk evidence and implications for developing-country agriculture. The first of a planned series of studies assessing the changing risk landscape affecting African agriculture is reported by Hurley, Pardey and Koo (2010).

pest and disease occurrences that are typically targeted to more localized (farm-level) and tactical (within season) crop management decisions.

2. Simulating the Spatial Occurrence of Crop Pests with CLIMEX

Climate is the primary driver of biological processes in different regions and seasons (Andreawartha and Birch 1954; Woodward 1987). This fact enables us to link global climatic data with observations on the geographical distribution and seasonal pattern of growth and stress of most poikilothermic species (plants, pathogens and cold-blooded animals). In this way we can infer the *potential* for population growth and decline and establishment of a species at any location, in the absence of constraints not related to climate, such as dispersal limitations and host presence. Such capacity is needed when assessing risks to agriculture or natural environments from pests, diseases and weeds under past, present or future climates.

Pests and pathogens are global problems that require global and regional risk assessments. Any model that is used to assess those risks must be capable of providing reliable results for any place on earth. CLIMEX (CLIMatic indEX) is a simplified computer model that infers the response of a species or other biological entity to climate from its geographical distribution and seasonal patterns of growth and mortality in different locations.[3] It was designed to extract the maximum amount of information from field observations and experimental data when there are insufficient data or necessity to build a detailed simulation model of the species. More specifically, the aim is to define the climatic regions that a species could potentially occupy in the absence of all other constraints such as mobility, predation, soil types, availability of hosts, and so on. In the context of biosecurity and invasive species, CLIMEX defines the *area at risk* of invasion by foreign pests, diseases and weeds should they be introduced from one continent to another. It also delineates the geographical extent and likelihood of establishment and persistence of crop pathogens in the presence of susceptible crop varieties and absent crop management practices designed to avoid or mitigate the presence of the pathogen. Within these ranges, CLIMEX defines the potential for seasonal population growth and survival at any given location. In this way it creates a seasonal and geographical profile of the effect of seasonal climatic variations on a species (that is, the species' phenology). Hence it is most relevant to the much larger scale of regions or continents than are field-scale pathogen models, which require site-specific data that are typically not available on regional scales.

[3] The initial version of CLIMEX was developed by Robert Sutherst and Gunter Maywald in 1985 under the auspices of CSIRO's Division of Entomology and has undergone numerous enhancements and upgrades. CLIMEX is now a particular implementation of DYMEX, a more general dynamic modeling platform.

2.1 Other Approaches and Applications

Other Pest Models

On April 6, 2007 HarvestChoice convened an expert consultation in St Paul, "Modeling and Evaluation of Biotic Constraints," to solicit expert opinion on pest modeling methods and procedures. While there are several global-scale, geospatial pest and disease models available, the consensus of the experts in attendance was that the CLIMEX method and software was most appropriate to pursue the goals of the HarvestChoice effort. In the following paragraphs, we briefly summarize some of the other available approaches.

There are many published studies that use logistic regression or related techniques to account statistically for the observed presence or absence of a particular species at a particular locale.[4] An array of geo-referenced historical climate variables (e.g., mean rainfall, monthly climate extremes and elevation) are used as independent variables to account for the occurrence (or otherwise) of the species.[5] The estimated coefficients are used in conjunction with observed, geo-referenced climate data to predict the probability of occurrence over landscapes beyond the range of the observed pest occurrence data. Like all regression models, the goodness of fit is optimized around the mean of the observations. Thus, while statistical climate matching models of this type may have some merit as a means of interpolating between sample data to fill-in missing occurrence data, Sutherst and Bourne (2008, p. 1235) conclude that, "… they are not appropriate for extrapolating beyond the data sets as is necessary with species invasion or climate change scenarios."[6] For example, consider the present application, where the intent is to assess the potential spatial occurrence of key crop pests, and suppose the goal were to determine the potential areal extent of the new *Ug99* variant of wheat stem rust (*Puccinia graminis* f. sp. *Tritici*). Use of a fitted statistical relationship between climate and the *observed* occurrence of *Ug99* would, by dint of the limited record of confirmed occurrences of *Ug99*, grossly underestimate (and thereby seriously misrepresent) the potential areal extent of this specific pathogen variant. Such spatial prediction error could be especially egregious if resistant wheat varieties or crop management methods that curtail the spread of this disease are not forthcoming.

Other systems are more similar to CLIMEX in their approach. For example, the NCSU-APHIS Plant Pest Forecast System (NAPPFAST) is equipped to utilize information about pest

---

[4] These and similar models are variously called ecological niche models, climate matching models, or rule-based approaches. See, for example, Guisan and Zimmerman (2000), Kriticos and Randall (2001), and the references therein for a list of such studies. These approaches are often used (questionably in some if not many cases) to project changes in the spatial occurrence of species resulting from climate change or species invasion events (e.g., Thuiller 2003).

[5] In a study supported by HarvestChoice, Hererra (2010) provides a comparison of several popular "ecological niche" models applying various models to generate pest risk maps for cassava pests.

[6] See also Randin et al. (2006).

biology along with gridded weather data to predict pest occurrence. The NAPPFAST system is composed of proprietary weather data, a system of templates for simplified model creation and a climate matching system that infers parameters from observed pest distributions (Magarey 2007).[7] There are some notable differences between CLIMEX and NAPPFAST. For example, NAPPFAST uses daily weather data whereas CLIMEX in its typical configuration uses 30 year averages of weekly data. The long-run average data used in CLIMEX are more appropriately called "climate data" as the effect of individual weather events is muted.[8] Further, the entire NAPPFAST interface is online (www.nappfast.org) whereas the most commonly used implementation of the CLIMEX model is a software package by the same name. Most importantly, CLIMEX enables modelers to easily use both deductive and inductive modeling techniques in an iterative fashion to fit the model.

Occurrence Mapping

There are a number of prior and some other on-going efforts to develop regional or global maps of the spatial occurrence of crop pests and diseases. Perhaps the most widely known initiative is that of CABI, which produces the various pest maps contained in the various CABI *Crop Protection Compendia*. Figure 1 is an example of a CABI pathogen map, this one indicating global distribution of wheat stripe rust (caused by *Puccinia striiformis* f.sp. *tritici*). The maps represent a compilation of historical crop pest occurrence information compiled from the published scientific literature. Dots on the map represent reported historical occurrence by country. The dots themselves are a binary indicator of occurrence within the surrounding country boundary, and otherwise have no geo-referenced relevance. Nor do they meaningfully indicate the geographical extent, timing or intensity of the pathogen occurrence within each country.[9] While the published pathogen literature compiled in the CABI Compendium is an invaluable resource, the maps provide little added analytical or operational value beyond visualizing the past occurrence of each pathogen on a country-by-country basis.

AgroAtlas (agroatlas.ru), a joint venture launched in 2003 between the USDA and three Russian scientific agencies based in St Petersburg, has developed a large number of geo-referenced maps indicating the potential extent and frequency of occurrence (differentiated into high, moderate and low classes) of a large number of crop pests, diseases and weeds for Russia and neighboring countries. Each map represents a fusion of information from published literature and expert judgment, with the area of distribution in some instances demarked by

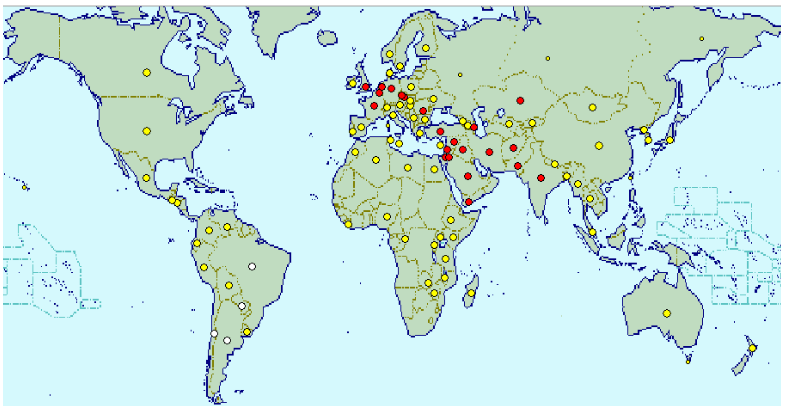
[7] NAPPFAST was developed in a joint venture involving the USDA's Animal and Plant Health Inspection Service (APHIS), North Carolina State University (NCSU) and the information technology company ZedX Inc.

[8] HarvestChoice is working to develop more options for weather input to enable time and location sensitive analyses.

[9] The color of the dots indicates that the pest is "present, no further details" (yellow), "widespread" (red), or "localized/occasional" (white).

the largest geographical (historical) extent of the crop in question. Although the maps are not generated by formally modeling the phenology of each pathogen, there is, presumably, an informal attempt to link historical (on average) climate information to the biology of each pathogen in a spatially explicit framework.

Figure 1: *Puccinia striiformis occurrence*



Source: CABI (2006).

The Global Biodiversity Information Facility (GBIF) is a distributed database that provides a portal to access museum data from around the globe (see www.gbif.org). In most cases, the records available from the data portal have been geocoded. The spatial resolution of the geocoded records varies. Data coverage is patchy, depending upon which specimen collection agencies are contributing to the GBIF framework. This dataset includes numerous geocoding errors, and includes records georeferenced to country centroids, without being tagged as such. Nonetheless, this data portal can be a valuable source of information for developing pest maps and potential distribution models, though one must exercise caution when cleaning and preparing the data for such purposes.

A large number of regional and local studies provide a wealth of pest occurrence data.[10] These studies are utilized by CABI (see above) and others to generate historical occurrence maps. The ever-expanding pool of local and regional studies are summarized by many researchers, perhaps most notably by Oerke et al. (1994) and Hill (1983 and 1987). While the data gleaned from these local and regional studies provide a wealth of information, they are not well-suited to the requirements of HarvestChoice. Most importantly, the various datasets from the studies are not standardized, such that experimental data are mixed with field observations, spatial scales are variable (or sometimes unknown), and even the definitions of intensity and occurrence differ across studies. As a result, these data are not well suited to developing consistent and reliable global-scale occurrence datasets.

[10] These include studies that only report occurrence among an ancillary set of variables where the primary objective was to measure other variables, for example crop variety yield trials where pest occurrences and other site and study attributes may (or may not) be systematically recorded.

Tracking and Monitoring

There are numerous examples of efforts to provide publicly accessible scouting and tracking information on specific crop pests and diseases in real (or near real) time. One of the more recent and prominent efforts is the RustMapper application jointly managed by the FAO Rust Spore Global Wheat Rust Monitoring System and the Borlaug Global Rust Initiative and hosted at CIMMYT (see www.fao.org/agriculture/crops/rust/stem/stem-rustmapper/en/). The initiative conducts regular geo-referenced field surveys of the occurrence of stem rust, testing for the presence of Ug99 variants, and sharing that information via a Google Earth enabled platform in conjunction with a host of other related information.[11]

Several other similar examples exist, for example:

- The Integrated Pest Management Pest Platform for Extension and Education (IpmPIPE) includes scouting information on a number of legume insects and pathogens in addition to tracking southern maize rust, cucurbit mildew and the pecan nut casebearer (see www.ipmpipe.org). The effort relies on field observations made by a network of cooperators at land-grant universities along with frequently updated state-by-state briefings from experts. IpmPIPE consolidates effort, data and funding from numerous organizations.
- The soybean rust (and aphid) tracking system managed by the Southern Region Integrated Pest Management Center located at North Carolina State University (see sbr.ipmpipe.org/cgi-bin/sbr/public.cgi).
- The National Invasive Pest Initiative in Australia has developed a pest alert system based on reports from crop consultants, which are collated and disseminated (see www.csiro.au/partnerships/NIPI.html).
- The Australian Plague Locust Commission undertakes extensive monitoring of locust populations and uses population dynamics and spread models to provide spatially-explicit pest alerts. These alerts have been used for area-wide management of locusts (see www.daff.gov.au/animal-plant-health/locusts).
- PestFax 2010 is a new pest alert system serving Western Australia. The service both receives and distributes reports of pests that are of importance to crops and pastures within the WA grain belt (see www.agric.wa.gov.au/PC_93358.html).

Some of these undertakings track observed pest and disease occurrence based on scouting data, others are linked to efforts to forecast the occurrence of crop pests and diseases and issue geo-referenced risk assessments.

[11] In 2009 surveys were conducted in Afghanistan, Azerbaijan, Eretria, Ethiopia, Georgia, Italy, Kenya, Saudi Arabia, Syria, Sudan, Tajikistan, Uzbekistan, Yemen, Zambia, and Zimbabwe. In 2010, surveys completed to June 2010 included, Azerbaijan, Egypt, Kenya, Iraq, Sudan and Syria.

(Risk) Forecasting

Properly forecasting pest risks is a demanding and information-intensive undertaking, especially when results are based on short-run weather data or are to be used to support immediate, farm-level decision-making. In short-run models, pest occurrence is predicted using weather forecasts and some limited knowledge of how the pest reacts to various weather events. Absent scouting and monitoring data, it is not possible to know whether a pest is present at a certain location at any given time. Thus, those systems that combine scouting with short-run predictive models may provide actionable information to those charged with responding to pest events. Systems that rely only on modeled pest responses and short-run weather data can only provide short-run potential occurrence maps which may assume (often implicitly) that the pest is actually present to begin with. That is, the systems that are not based on scouting cannot warn farmers of an impending pest event, only that conditions are favorable (or not) for a pest event to occur. Farmers may respond to such information by employing prophylactic measures, but without actual evidence of nearby occurrence it may often be uneconomical to take action. For example, such models will likely highlight entire regions as being at risk of infection even when the pest is not present in that region.

A few examples of existing risk forecasting applications include:

- The US. Fusarium Head Blight (head scab) Risk Assessment Tool created and maintained by a consortium of universities[12] and funded by USDA-ARS (available at www.wheatscab.psu.edu/riskTool_2010.html).
- The PA PIPE effort is an example of a state-level forecasting system (available at extension.psu.edu/pa-pipe). The PA PIPE is a joint effort of Penn State and ZedX and is apparently not part of the IpmPIPE system. PA PIPE provides modeled results for about 35 weed, insect and disease pests of a variety of crops, along with other climate and crop phenology information.
- Rutgers University and the New Jersey Cooperative Extension maintain a statewide network of insect traps. European Corn Borer and Corn Earworm counts for each trap are collected weekly, mapped and reported via the internet and fax (see www.pestmanagement.rutgers.edu/IPM/Vegetable/).
- The aWhere Pest Monitor provides risk predictions for several plant diseases in south and eastern Sub-Saharan Aftica. Predictions are given up to nine days in advance. While the service is presented as a “monitor,” it is not clear whether scouting data are used in making the predictions (see www.awhere.com/CSISA/Solutions/pestmonitor.aspx).

[12] The US Fusarium Head Blight Risk Assessment Tool is maintained and developed by the following universities: Penn State, Ohio State, Kansas State, Purdue, North Dakota and South Dakota.

- The Penn State University's PESTWATCH system provides real time and historical scouting reports for several sweet corn pests. The scouting sites that report to this system are located in several states, but are most heavily concentrated in Pennsylvania, New York and Deleware (see www.pestwatch.psu.edu/sweetcorn/tool/tool.html).
- The Department of Agriculture and Food in Western Australian provides modeled predictions for a variety of pests (see www.agric.wa.gov.au/PC_92989.html). Different models are used for the various pests, for example the barley yellow dwarf virus(BYDV) prediction system described by Thackray, Diggle and Jones (2009).

Strategic vs Tactical Decision Making

Many management scientists make a clear distinction between strategic and tactical decision-making.[13] Tactical decisions are generally aimed at *responding* to events that are either occurring, or that are expected to occur in the near-term. By contrast, strategic decisions are intended to influence the environment in which events occur. Each type of decision-making is made possible by different types of information.

Tactical decision-making processes are supported by data and information that describe what is actually happening, or what is expected to happen soon. In the present context, most of the pest monitoring and prediction systems support pest management decisions over days or perhaps the current season. That is, the decisions are implemented to r*espond to* a current problem. While it is often important to respond to pest events, such responses may be second-best solutions.

Strategic decisions aim to influence the very set of situations that will be encountered (and responded to) in the future and, as a result, are less relevant for current situations. Rather, strategic decisions are founded on information that can be used to answer hypothetical or abstract questions such as:

- What would be the consequences of removing a given pest at a certain location? or,
- If a new variety were planted in a previously un-cropped area, what pest risks might the newly planted fields face?

Strategic decisions may well change fundamental aspects about the future, such as the location of crop production or the crop management and crop varietal options resulting from investments in R&D, all of which can have substantial consequences for crop pest problems in the years ahead. Exclusive reliance on observed pest occurrence data provides only a partial picture of the plausible scenarios on which strategic research investment and other crop production choices must rely.

[13] There is also an "operational" decision-making category, which includes the myriad day-to-day decisions that must be made in any enterprise. These decisions affect immediate or very short-run outcomes, but may also have longer term consequences.

3. Mapping Crops Pests and Diseases

3.1 CLIMEX Metrics of Spatial Suitability

The CLIMEX model has two primary outputs: a Growth Index and an Ecoclimatic Index:

- The annual Growth Index GI_A is a summary of the potential for a given pest to *thrive at a location*, absent any stressors (heat, cold, wet, dry). Conceptually, the annual and weekly Growth Indices summarize the species' growth response to temperature and moisture (but can include other factors that are needed for growth such as radiation and daylight).
- The Ecoclimatic Index represents the ability of a species to *persist* at a given location. Ecoclimatic Index values combine a species' growth potential with its reaction to various stressors, primarily extreme hot-,cold-,wet- and dry-stresses and their interactions (e.g., hot-wet stress).

Both indices refer to the species population responses to local climate conditions and the pest's phenology, absent any other constraints (i.e., assuming there is a suitable host plant present).[14] Ecoclimatic Index values are calculated as the product of the annual Growth Index and indices of the species' response to the various stressors. To the user, stresses are scaled from 0 (no stress) to infinity, with a value of 100 being lethal. For the calculation of the Ecoclimatic Index, each stress factor is rescaled from zero (lethal stress) to one (no stress) so that their product has the same range.[15] Therefore, the annual Growth Index value is an upper bound for the Ecoclimatic Index value at a location. Further, if the combined stress indices are not lethal, the species can potentially establish so long as the necessary growth conditions are satisfied.

3.2 Inferential Mapping with CLIMEX

CLIMEX can be used in several modes to explore different questions or to take advantage of different sources and types of information on a species' ecology. An implicit assumption is that in order to persist at a location, a species must be able to tolerate an inclement season, and during the favorable season it must be able to grow sufficiently to reproduce. CLIMEX is based on the premise that the *potential* geographical distribution of a species is the end product of

[14] For some of our subsequent analyses, it is critical that we model the spatial distribution of the pest separately from (or unconstrained by) the present distribution of host plants. For example, we plan to assess the susceptibility of the world's wheat crop to stem rust if the location of wheat production should shift over time for, say, economic reasons. That is, we remove the "host" leg of the disease triangle (see Section 4).

[15] Conceptually, the stress indices are reversed before taking the product so that a value of one indicates no stress. When multiplied together, the stress indices will equal zero if *any* of the stressors are lethal. This is convenient since it allows growth index values to readily be recalibrated to form the Ecoclimatic index and for a single stressor to limit establishment.

the species' responses to climate.[16] Given this assumption, each map produced by CLIMEX represents the species' estimated response to factors such as temperature, moisture and day length within a region. The maps therefore represent the set of assumptions regarding the species responses to each of the climatic variables and their extreme values, separately and in combination at every location within that region.

CLIMEX can be used inferentially to estimate a species stress responses based upon its known distribution. If direct experimental observations of a species growth response to climatic variables are available, it can also be used deductively to project its growth response to different climates. Some growth parameters can also be inferred by matching phenological observations and the corresponding seasonal climatic variables. In practice, for any species modeling exercise, a range of techniques and approaches are employed to address a range of different questions and deal with a range of primary data constraints.

The boundaries of a species' potential distribution are determined by the availability of water, energy and other factors needed for growth along with the duration and severity of each of the stresses. By contrast, at any given time, the *actual* distribution of a species may be further circumscribed by exogenous constraints, such as the availability of hosts, excessive competition or predation. Consequently the potential spatial distribution of a pest derived from a CLIMEX model may extend well beyond the present or historical extent of the primary crop that is targeted by the pest (e.g., Kriticos et al. 2007; Watt et al. 2009). This is a desirable attribute from the perspective of *HarvestChoice*, where the intent is to assess the production losses or potential cost of control "with" versus "without" the pest. Relying only on observed occurrences, one may grossly misrepresent (and, likely, underestimate) the potential crop productivity losses or increased costs of production incurred as a consequence of the pathogen. The observed occurrence may be (heavily) circumscribed by the present location of production or preemptive efforts to avoid or mitigate the presence of the pest via the use of resistant varieties or crop sprays. In practical terms, assessment of a crop technology that might enable production in environments now thought to be inhospitable to the plant requires that the potential pest and disease risks in the new locations be understood. Limiting assessment to areas in which production already takes place would ignore exactly the types of information that models such as CLIMEX help reveal.

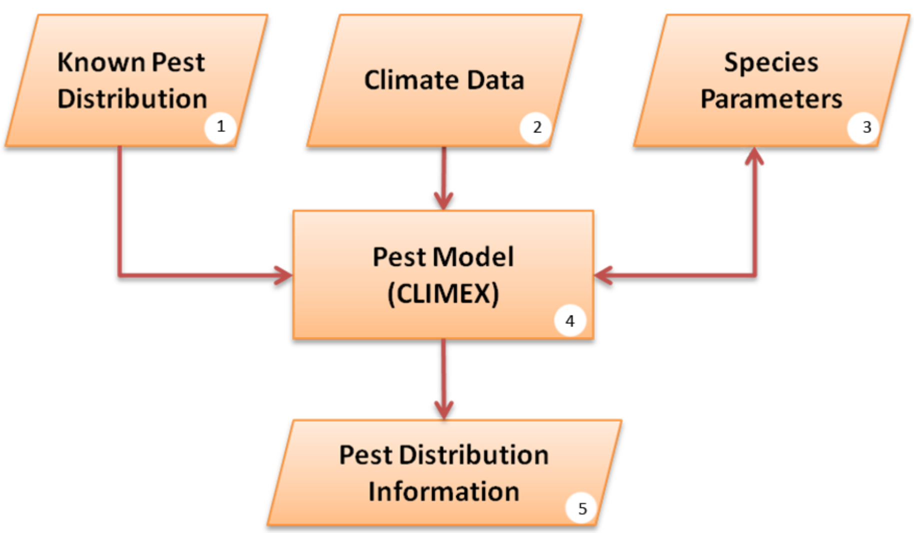
The extent to which a species is able to withstand climate stresses depends partly on the size of the population that is produced during the growth season. Hence, a large value of the annual Growth Index represents an increased capability of the population to persist

[16] Clearly the observed or actual spatial occurrence of a particular pathogen is affected by climate attributes that are not accounted for by the CLIMEX model, such as wind direction and speed, humidity, soil types, human mediated movement, presence of primary and alternate hosts, and host plant resistance. Here the goal is to map the potential distribution of the pathogen assuming temperature, moisture, daylength attributes and solar radiation are the determining factors.

through the stressful season. To properly capture these processes we need to map both the Growth Index and each of the stress variables, and then map the combined growth and stress response in the form of the Ecoclimatic Index. This is in contrast to statistical models used to map presence or absence data.

Using CLIMEX, with appropriate geo-referenced climate data in hand, there are two options to developing potential pest distribution maps (Figure 2). One approach is to use reported pest occurrence data spanning a suitably diverse ecological transect (Figure 2, Object 1) to infer the appropriate species parameters (Object 3) by iteratively fitting the CLIMEX derived species occurrence map to the reported spatial distribution. Another approach is to start with known species parameters (delimiting critical thresholds for temperature, soil moisture, and other variables that define the growth and stress profiles of the pathogen (Object 3), and use CLIMEX to generate the appropriate occurrence maps (Object 5). In practice both methods are often used in an iterative fashion: i.e., beginning with plausible species parameters (see Box 1) that accord with the reported scientific literature or the considered judgment of knowledgeable individuals about the phenology of the pest, to develop occurrence maps that are spatially consistent with the available observational evidence, and vice versa.[17]

Figure 2: *Schema for Inferential Pathogen Occurrence Mapping using CLIMEX*



Source: Authors.

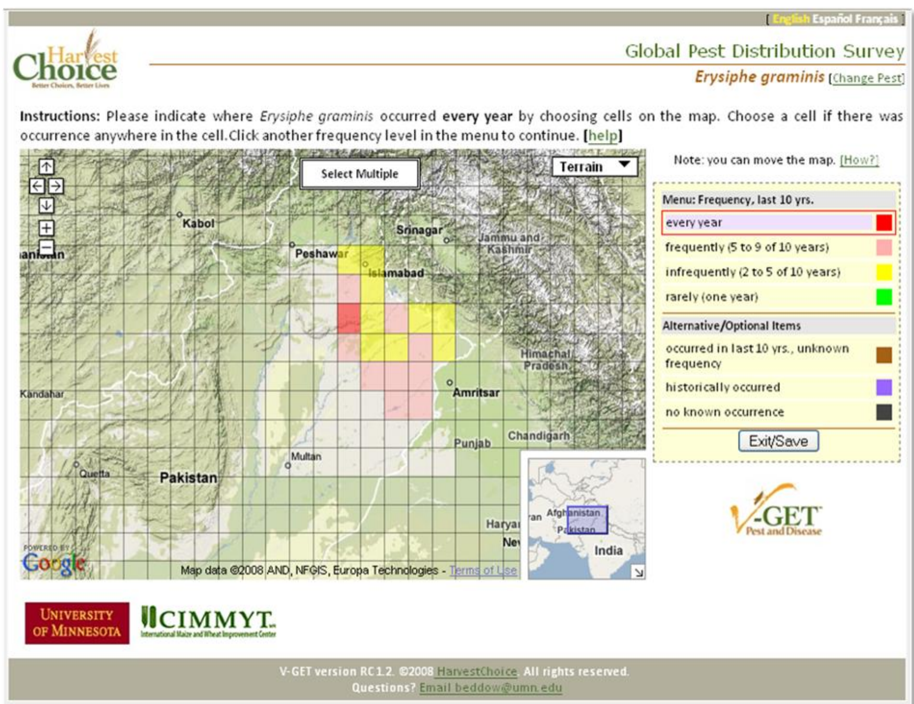
V-GET™ (Virtual Georeferenced Elicitation Tool)

Unfortunately, published information on the spatial distribution of important crop pests is sketchy: either the available information is sparse or it lacks sufficient geographical specificity

[17] CLIMEX cannot be used to deterministically or statistically derive species parameters from known species distributions. Rather, available data are used to infer species parameters in consultation with experts who are able to determine the veracity of the inferred phenological response.

to be valuable for modeling. However, it seemed there was a large, hitherto largely untapped reservoir of tacit knowledge on the spatial occurrence of crop pests that could be used to calibrate a CLIMEX model. To capture that tacit knowledge the HarvestChoice team at the University of Minnesota developed V-GET™, a web-enabled geo-referenced survey tool. For this particular application, the tool facilitates a web-enabled survey of spatially explicit data on the location and frequency of occurrence data for specific pests and diseases from geographically dispersed respondents. V-GET™ is built around the Google Maps application programming interface (API), which allows for manipulation and distribution of mapped data via Google's servers while overlaying a structured survey. Respondents can provide spatial data by clicking a displayed map using an intuitive interface (Figure 3).

Figure 3: *V-GET™ Data Capture Screen for Crop Pest Survey*

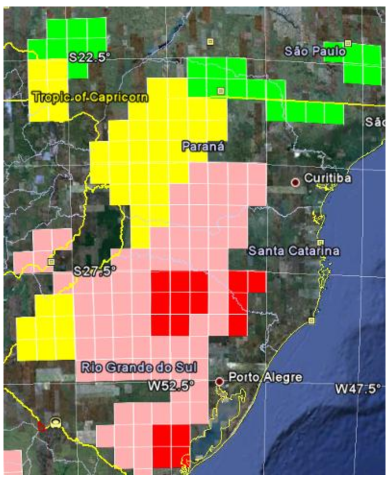


Source: Authors

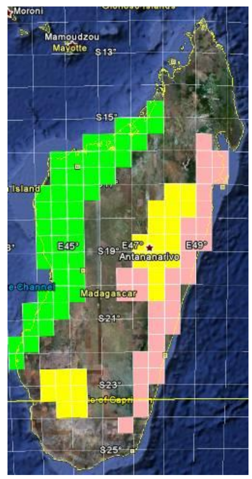
V-GET™ allows for such complexities as multiple geographically dispersed survey administrators, and elicitation of information from each respondent on multiple sub-topics (e.g., different pests). In this instance respondents were identified by key HarvestChoice collaborators based on their scientific knowledge and field experience with the pathogen and crop in question. Figure 4 provides an example of the responses received on the known distribution of several pathogens from several respondents. Information from multiple respondents for each of the pathogens surveyed was consolidated to form a composite response. This was coupled with information gleaned from a systematic search of the published literature to calibrate the CLIMEX model. Where feasible the observed (i.e., solicited or published) occurrence information was partitioned into two data sets. Observations from one particular region were deemed in-sample data and used to fit a simulated map and refine the various pest parameters. Then occurrence data from other regions were treated as out-of-sample information and used to validate the map by assessing the spatial concordance of the out-of-sample observations with the CLIMEX extrapolations based on the inferred species parameters obtained using the in-sample data.

Figure 4: *Examples of V-GET Survey Responses*

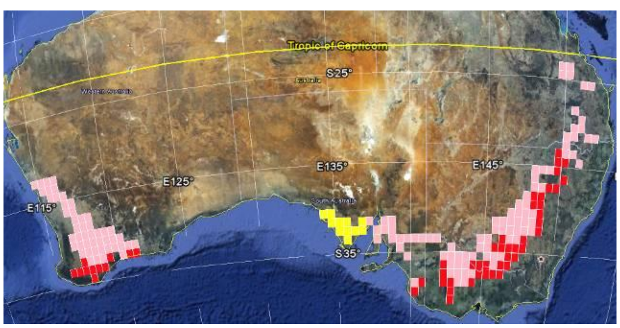
Panel a: *Fusarium graminearum* (scab)



Panel b: *Magnaporthe oryzae* (rice blast)



Panel c: *Puccinia striiformis* (wheat stripe rust)



Source: Occurrence data are from one or more V-GET survey respondents per figure and the underlying map data are from Google Earth.[18]

Inferential Mapping Steps

It would be difficult to fully understand the HarvestChoice pest and disease maps without some knowledge of how they were produced. Here we provide an abridged (and somewhat stylized)

[18] The V-GET tool also provides output in formats that can be easily used in ArcGIS and other mapping software and analytical packages. The "quick view" feature demonstrated here allows survey administrators to quickly and easily map results from a selection of respondents using Google Earth.

set of basic steps that HarvestChoice researchers employ to derive spatial data on potential pest and disease occurrence.

Step 1−Baseline Climate and Species Data: We do not begin by modeling global pest occurrence. Rather, we first choose one or more regions and attempt to fit modeled occurrence within that region. The starting point for any risk analysis of a new species is to acquire a sound knowledge of the geography and climate of the region of interest. For example, the gradients of temperature and moisture across the landscape must be identified. The climate data used in our CLIMEX modeling effort consists of a 30 year average (1970-2000) of weekly observations reported on a 30 arcminute global grid.[19] The source data were assembled by the Climate Research Unit (CRU) of the University of East Anglia.[20] In parallel with constructing mapped climate gradients, it is necessary to 'see the world through the eyes of the species being modeled' and understand how the species responds to each of the primary climatic variables of temperature and soil moisture. Armed with this knowledge, the model then allows one to spatially visualize how the pest will respond to the different climates it may encounter. Thus, the climate database and literature survey of the species of interest provides the wherewithal to interrogate the climate data to infer how the species responds to the temperature, available moisture (in the form of soil moisture) and other factors at each location. Information on the known spatial distribution and seasonal patterns of development of each species are gleaned from comprehensive searches of the literature and extensive consultation with experts.

Step 2−Inferring Species Stress Parameters: Initially using the native range of the species, this discovery process involves iteration of the parameter values for each of the stresses, preferably until model estimates of the potential geographical distribution of a pest (its Core Distribution) agrees with the observed distribution in one or more of the in-sample regions. Each stress variable has two parameters that represent the stress threshold value and the rate at which the species accumulates stress, respectively (see Box 1). Once these parameters are estimated, the model results are examined to determine how and where in the modeled region the different stresses are experienced by the species. It is frequently observed that species can experience a climatic range expansion when released from the effects of their natural enemies (Kriticos and Randall 2001; Keane and Crawley 2002; and Sutherst and Bourne 2009). This observation led to the recommendation that when modeling invasive species, exotic range data should be included wherever possible (Kriticos and Randall 2001). The implicit assumption in purely inferential modeling is that the species has reached its potential limits. Without corroborating evidence from another knowledge domain such as ecophysiological

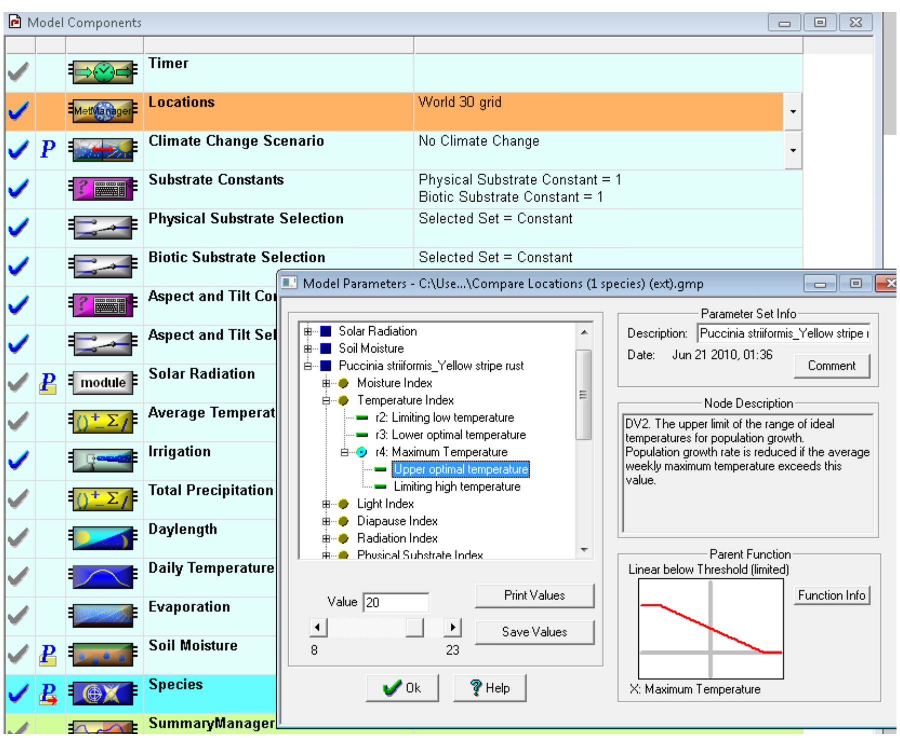
[19] A 30 arcminute (30′) resolution generates cells that are about 55×55 km at the equator.

[20] The CRU data can be downloaded from www.cru.uea.ac.uk/cru/data/. For mapping at 30′ resolution, the climate data we used were pre-formatted for CLIMEX and were included with the CLIMEX version 3 software. Some maps were created at 10′ resolution, the data for which was downloaded from CRU in 2003.

Box 1: CLIMEX Variables

The CLIMEX model includes a number of parameters that are used to define the climatic growth response of a species. The model encapsulates an assumption that growth occurs within a specific range of environmental conditions. While, for at least a limited time, it is possible for a species to tolerate climatic conditions outside the growth-enabling ranges, certain extreme conditions can cause stress, which can accumulate to reduce or eliminate a population. Thus, the parameters can be coarsely grouped into two categories: growth parameters and stress parameters. Some parameters are used to define a simple distribution of the organism's growth response to a climatic variable. For example, a species-specific function describing the growth response to temperature can be defined by specifying four variables: the temperature below which growth cannot occur, the lowest and highest temperatures for optimal growth, and the maximum temperature beyond which growth cannot occur. Stress response functions are handled differently and, at a minimum, include parameters to specify the value at which stress begins to occur and the rate of stress accumulation after that point. For example, the dry stress response is specified as a linear function beginning at a threshold moisture level and increasing at a specified rate as the duration of the stress experience increases. More detailed information on the CLIMEX parameters (and the model in general) can be found in the CLIMEX User's Guide (available from www.hearne.com.au).

Figure 5: Specifying a CLIMEX Model



experimentation, this assumption should always be treated with some suspicion. This is especially true when an invasive species has recently arrived.

Step 3−Estimating Spatial Metrics of Potential Occurrence: Once the stresses have been estimated from the edges of the geographical distribution (if available), the seasonal phenology of the species in different regions is used to estimate the growth parameters. Recall that the weekly Growth Index GI_W simulates the seasonal pattern of potential for growth of the species. To form the Growth Index, CLIMEX takes weekly snapshots of the climatic potential for growth and not the pest population during the season. High GI_W values indicate the weekly climatic potential for high reproduction and organic growth rates (see Box 1).

Step 4−Mapping Pest Occurrence: When both the stress and growth variables have been optimized it is possible to examine the estimated combined response of the species to both favorable and stressful values of each climatic variable, in the form of the Ecoclimatic Index. This gives a single number for each location that provides a high-level summary of the species' performance at each location (see Box 1).

Step 5−Model Validation: Ideally, once verified, models should be validated using independent observations that are unrelated in any way to those used to estimate the parameter values. In other words they are typically from another region and they are best kept 'blind' until used for the validation. Once formulated, the model represents the hypothesized response of the species to climate. The model can be refined progressively as more information becomes available. There may be several reasons why a model cannot be validated, including an absence of sufficient distribution data on a continent not used for model calibration. A frequently overlooked option is to use data from a completely different knowledge domain to validate different aspects of the model. For example, phenological observations may be used to validate the growth parameters, which may have been derived from direct experimental observations. In practice, it is unusual for a model of such complexity as CLIMEX to be comprehensively validated. It falls to the model authors to describe their model, and the extent to which it might be reliable, drawing attention to areas of particular uncertainty. Sensitivity analyses can inform this reliability (e.g., Venette and Cohen 2006).

Step 6—Revisions: After validating a pest map, HarvestChoice assigns the map a version number and submits it for peer review by leading experts on the species. Depending on the comments returned from the reviewers, the map (and its underlying model parameters) may be revised and resubmitted for review. Once the map is found to be acceptable, it is released for public use and review, but the map is never considered final. Rather, the maps are open to continual refinement and revision as new information becomes available.

4. Interpreting CLIMEX Pest Occurrence Maps

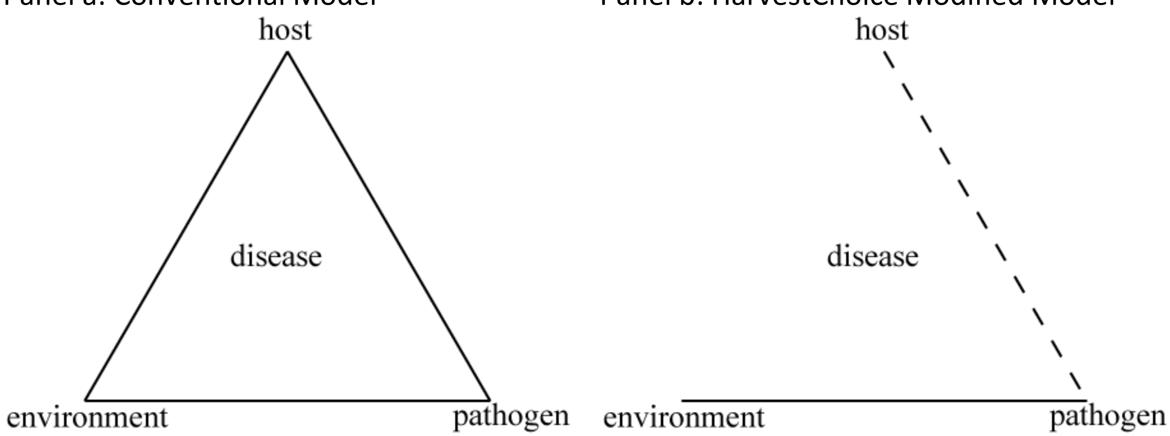
The Disease Triangle

A core concept in plant pathology is that the study of plant diseases must consider the interactions between host plants, the environment and pathogens. This idea is often taught in introductory plant pathology courses by presenting a “disease triangle” (Figure 6, Panel a). Among other things, the triangle emphasizes that a pathogen cannot cause disease without a susceptible host and a favorable environment, and a host may be more or less susceptible depending on its environment. Wherever we have a report of a disease, we can be sure that the disease triangle is complete and sufficient for the host and the disease. However, we should be mindful with crop pests in particular that the climatic environment could be strongly and significantly modified by human cultural practices such as irrigation. In CLIMEX we can explore this potential using irrigation or climate scenarios. If we use all of the distribution records that we are confident are based on dryland conditions, and we use CLIMEX to apply the resulting fitted model to a climate dataset, then in the first instance we are producing a map where the disease triangle is conceptually reduced to the representation shown in Figure 6, Panel b.

Figure 6: *The Disease Triangle*

Panel a: Conventional Model

Panel b: HarvestChoice Modified Model



Source: The disease triangle (Panel a) was recreated from Schumann and D’Arcy (2006, p.4)

In fitting this model, we also need to be mindful of how we framed the problem. If we rely upon distribution data to inform the fitting of stress parameters, then we may be including parameters that describe how the host is limited by climate, rather than the pest *per se.* For example, a fitted model of *Dothistroma* needle blight was built in this manner, and many of the resulting stress parameters were related to the survival of the *Pinus* host trees (Watt et al. 2009). In that case, the presence of the host was considered secondarily to the climatic conditions suitable for disease expression on the host. Depending upon the specific system being modeled, there may be areas of underlap or overlap of host climatic suitability and pathogen climatic suitability.

The HarvestChoice pathogen maps indicate the *potential* for the pest to persist if the host were present. The disease triangle framework implicitly assumes that a susceptible host is available at every point on the globe, conceptually reducing the disease triangle to the representation shown in Figure 6, Panel b. The pathogen-environment connection is explicitly considered via the CLIMEX model as discussed in other sections of this document (represented by the solid segment in the diagram). The host-pathogen segment is dashed to emphasize that there is implicit consideration of the relationship, namely insofar as the environment affects infection, but the more intricate host-pathogen interactions are not considered. In CLIMEX, there is no direct connection between the host and the environment because the host is assumed to be present and susceptible to infection everywhere, regardless of the environment.

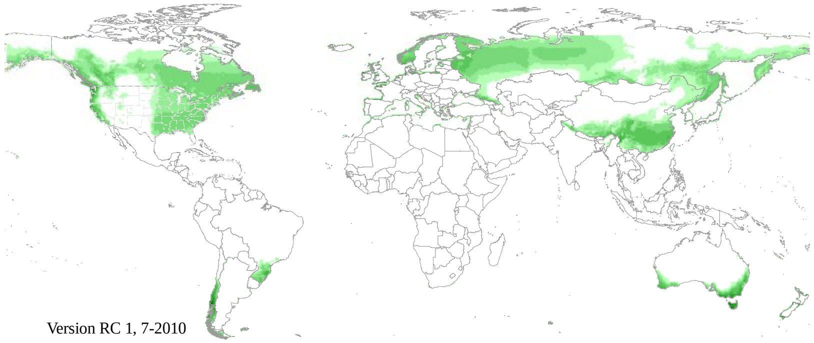
The conceptual disease model shown in Figure 6, Panel b cannot be used to generate maps of *actual* disease occurrence, even when given the assumption that the host is present everywhere. But, the assumption does allow us to consider a hypothetical case in which host plants are ubiquitous. Typical disease maps show an estimated actual distribution of the disease and are clearly useful. But, the conceptual model that enables creation of the actual occurrence maps also limits them. In such models, disease occurrence is limited to areas that are currently cropped. By assuming the presence of a susceptible host, the model can generate maps of *potential* disease distribution.

An Example: Puccinia striiformis

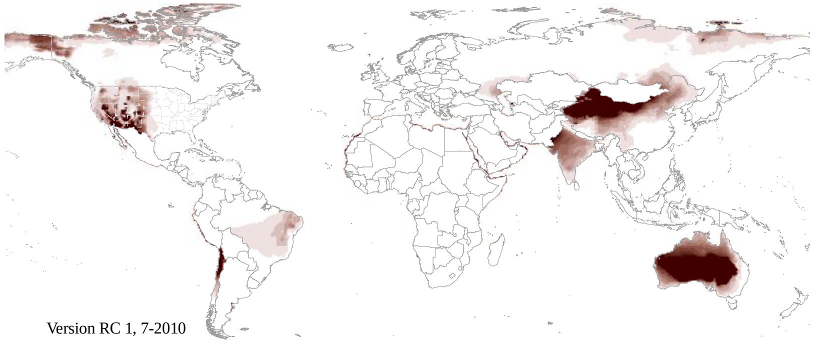
It is instructive to consider an actual example of CLIMEX-generated maps to highlight some of the features described above. Here we consider a version of the HarvestChoice potential distribution maps for *Puccinia striiformis*, a pathogen that causes stripe rust on wheat. Recall that the annual Growth Index is intended to show those areas in which water and temperature (among other factors) are within the ranges that permit the species to grow. Figure 7, Panel (a) shows the Growth Index for *P. striiformis*; the pathogen can potentially develop in the green shaded areas of the map, and darker shades indicate more favorable conditions. Those who have experience with *P. striiformis* may immediately object that the pathogen cannot develop in, say, Ontario north of the 49th parallel, because wheat is not widely grown there. However, the map makes more sense if one keeps in mind that it represents the risk in the event that a wheat crop (or another suitable host) were to be grown there. So, for example, if scientists were to develop a wheat variety that could enable commercial-scale production in these areas, one could use the map in Panel (a) to warn that the provinces' new wheat fields might encounter stripe rust problems. Such conclusions could not be drawn if the stripe rust distribution were modeled conditionally on the current wheat distribution.

Figure 7: *Growth, Stress and Persistence Indexes for Puccinia striiformis*

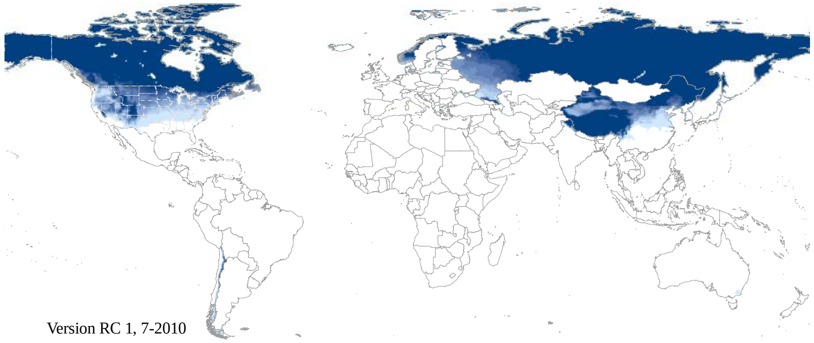
Panel a: Growth Index



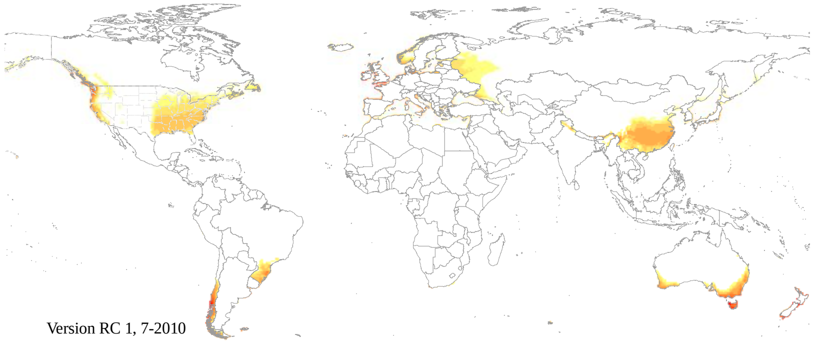
Panel b: Dry Stress



Panel c: Cold Stress



Panel d: Eco-Climatic Index



Source: Beddow et al. (2010a).

Maps for dry and cold stress are shown, respectively, in Panels (b) and (c) of Figure 7. These represent the extent to which each factor hinders survival of the species. For example, the cold stress map (shown in Panel c) highlights locations in which cold conditions limit the species' ability to survive the winter; values near 100 (indicated by very dark blue shades) imply that the pathogen will die out while lower values indicate potential for overwintering. Together, the stressors limit *persistence* in areas in which the pathogen can otherwise develop intra-seasonally. In the case of stripe rust, the stressors generally affect overwintering and therefore the potential to persist in the area.[21] Continuing the Ontario and Quebec example, note that provinces are colored in a very dark blue, indicating that the pathogen cannot survive a full season due to cold conditions. Of course, one needs to keep in mind the potential for the pathogen to be transported into any crops in these areas each year in time to damage the crop.

The Ecoclimatic Index summarizes the overall species response, factoring in both growth and stress (Figure 7, Panel d). Ecoclimatic Index values are indicators of the probability of inter-seasonal survival of the species, where high values indicate a high likelihood for survival. In the map, lightly shaded (yellow) areas have low index values, and dark (red) areas have higher values. Notice that the shaded areas in the Ecoclimatic Index map are a subset of the area shaded in the Growth Index map. This is because the species cannot establish in areas where it cannot develop; further, visual inspection of all four maps reveals that the potential persistence area is (roughly) the area shaded in Panel a, less the darkly shaded areas in Panels b and c.

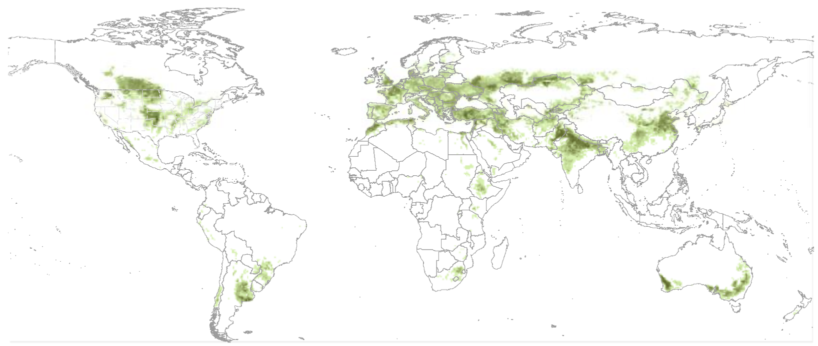
Overlaying Potential Pathogen Distributions on Cropped Area

HarvestChoice estimates of the year 2000 wheat growing areas are shown by the shaded areas in Figure 8, Panel a. To illustrate the strategic value of the potential pest distribution maps, the geo-referenced wheat geography information in Figure 8, Panel (a) was juxtaposed against the CLIMEX indices reported in Panels (a) and (d) of Figure 7 to provide an estimate of the potential range of growth (Panel b) and persistence (Panel c) of stripe rust within the year 2000 areal extent of wheat. This data overlay indicates that 37 percent of the world's wheat crop is vulnerable to infestations of stripe rust (meaning in this instance that they have Growth Indexes in excess of 20), while30 percent of world's wheat area has a high propensity to sustain stripe rust infestations year round (meaning they have Ecoclimatic Indexes in excess of 20) (Beddow et al. 2010b). Recall the climate data underlying these pest occurrence distributions is a 30 year average of weekly climate observations, so the share of global wheat acreage deemed susceptible to the disease is an indication of the potential for outbreaks and persistence of infestations "on average" rather than an estimate of the observed occurrence of the disease in any given year.

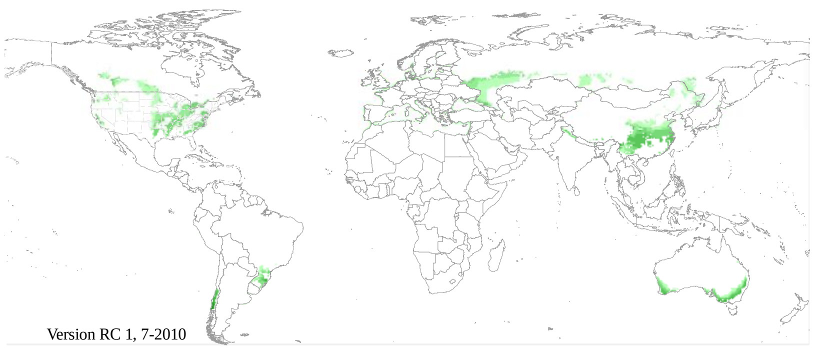
[21] Pests and pathogens can also face limits to over-summering. For example stem rust, which is more sensitive to hot-wet stress, cannot over-summer in hot, humid areas.

Figure 8: *Potential P. striiformis range, 2000*

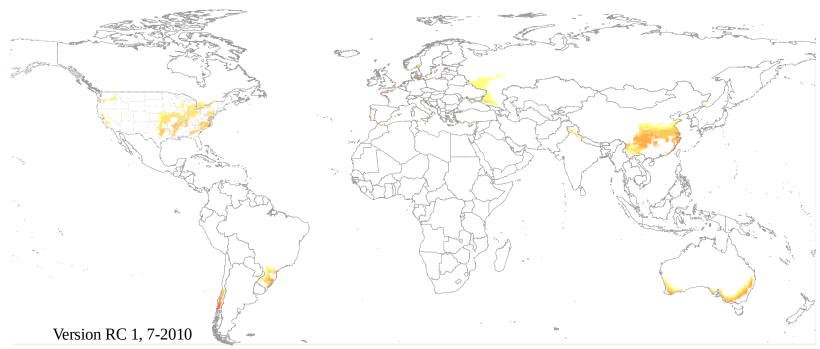
Panel a: Wheat Growing Areas, c. 2000



Panel b: Growth index for Wheat Areas



Panel c: Eco-climatic index for Wheat Areas



Sources: Panel (a) is from the HarvestChoice Spatial Allocation Model (SPAM) of You et al. (2000); Panels (b) and (c) are from Beddow et al. (2010a).

5. Conclusion

Crop pests are costly to control and result in widespread crop losses. But the exposure to and consequences of these pest problems do not impact all farmers in all locations equally, or at all times. Developing an informed perspective on the potential for pest outbreaks in particular locations (and, ideally, the frequency and intensity of such outbreaks) has value for a whole array of investments and actions designed to avoid or ameliorate the economic and social consequences of these crop pests.

Until now, there have been no comparable, geo-referenced estimates of the global potential occurrence and persistence of key crop pests to serve as a basis for setting strategic priorities to deal with these crop production problems. The pest mapping effort underway by *HarvestChoice,* and described briefly in this report, is part of a broader *HarvestChoice* undertaking to evaluate the potential crop productivity and production risk consequences of these biotic constraints. Research to develop new pesticides and insecticides or to breed new crop varieties that are resistant to these pests is costly, and it typically takes decades to realize useful results. Thus, informing decisions to invest in research and related strategies aimed at mitigating these pest problems can yield large and lasting returns, particularly if the limited research and agricultural development funding is directed to the areas with the largest economic and social payoffs. It is to this end that these new global pest occurrence data are being put.

References

Andrewartha H.G. and L.C. Birch. *The Distribution and Abundance of Animals*. Chicago: University of Chicago Press, 1954.

Beddow, J.B., E. Duveiller, P.G. Pardey and R.W. Sutherst. "Potential Global Distribution of Wheat Stripe Rust *(Puccinia striiformis)*." HarvestChoice Pest Brief. St. Paul and Washington D.C.: University of Minnesota and IFPRI, 2010a (in preparation).

Beddow, J.B., P.G. Pardey, E. Duveiller, and D. Hodson. "Wheat Rusts: A New Global Perspective." HarvestChoice Working Paper. St. Paul and Washington D.C.: University of Minnesota and IFPRI, 2010b (in preparation).

CABI. *CABI Crop Protection Compendium*. Data Disk. Wallingford, UK: CAB International, 2009.

Guisan A. and N.E. Zimmermann. "Predictive Habitat Distribution Models in Ecology." *Ecological Modelling* 135 (2000): 147-186.

Hill, D.S. *Agricultural Insect Pests of the Tropics and their Control*, 2 ed. New York: Cambridge University Press, 1983.

____. *Agricultural Insect Pests of Temperate Regions and their Control*. New York: Cambridge University Press, 1987.

Herrera, B.V. "Predicción de La Distribución Geográfica de Cuatro Especies de Patógenos de La Yuca a Nivel Mundial, Utilizando Modelos de Nicho Ecológico." M.S. Thesis. Cali, Colombia: Universidad del Valle, 2010.

Hurley, T.M., P.G. Pardey and J. Koo. "A Geo-Spatial Assessment of the Intrinsic Climate Risk Facing African Agriculture over the Past 100 Years." HarvestChoice Working Paper. St Paul and Washington D.C.: University of Minnesota and IFPRI: 2010 (in preparation).

Hurley, T.M. "A Review of Agricultural Production Risk in the Developing World." HarvestChoice Working Paper. St Paul: HarvestChoice, University of Minnesota, 2010. Available at www.harvestchoice.org/production/risk/intro.

Keane R.M. and M.J. Crawley. "Exotic Plant Invasions and the Enemy Release Hypothesis." *Trends in Ecology and Evolution* 17 (2002): 164-170.

Kriticos D.J., K.J.Potter, N. Alexander, A.R. Gibb and D.M. Suckling. "Using a Pheromone Lure Survey to Establish the Native and Potential Distribution of an Invasive Lepidopteran, Uraba lugens." *Journal of Applied Ecology* 44 (2007): 853-863.

Kriticos D.J. and R.P. Randall. "A Comparison of Systems to Analyse Potential Weed Distributions," in R.H. Groves, F.D. Panetta and J.G. Virtue, eds. *Weed Risk Assessment*. Melbourne: CSIRO Publishing, 2001.

Magarey R.D., G.A. Fowler, D.M. Borchert, T.B. Sutton, M. Colunga-Garcia and J.A. Simpson. "NAPPFAST: An Internet System for the Weather-Based Mapping of Plant Pathogens." *Plant Disease* 91 (2007): 336-345.

Oerke, E.C., H.W. Dehne, F. Schönbeck and A. Weber. *Crop Production and Crop Protection: Estimated Losses in Major Food and Cash Crops*. Amsterdam: Elsevier Science, 1994.

Randin C.F., T. Dirnböck, S. Dullinger, N.E. Zimmermann, M. Zappa and A. Guisan. "Are Niche-Based Species Distribution Models Transferable in Space?" *Journal of Biogeography* 33(2006): 1689-1703.

Schumann, G.L. and C.J. D'Arcy. *Essential Plant Pathology*. St. Paul, MN: The American Phytopathological Society, 2006.

Sutherst, R.W. and G.F. Maywald. "A Computerised System for Matching Climates in Ecology." *Agriculture Ecosystems and Environment* 13 (1985): 281-99.

Sutherst, R.W. and A. S. Bourne. "Modelling Non-Equilibrium Distributions of Invasive Species: A tale of two modelling paradigms." *BioIogical Invasions* 11 (June 2009): 1231-37.

Thackray D.J., A.J. Diggle and R.A.C. Jones. "BYDV PREDICTOR: a Simulation Model to Predict Aphid Arrival, Epidemics of Barley Yellow Dwarf Virus and Yield Losses in Wheat Crops in a Mediterranean-Type Environment." *Plant Pathology* 58 (2009): 186-202.

Thuiller W. "BIOMOD - optimizing predictions of species distributions and projecting potential future shifts under global change." *Global Change Biology* 9 (2003): 1353-1362.

Venette, R.C. and S.D. Cohen. "Potential Climatic Suitability for Establishment of *Phytophthora ramorum* within the Contiguous United States." *Forest Ecology and Management* 231 (2006): 18-26.

You, L, Z. Guo, W. Ojo, K. Sebastian, M.T. Tenorio, S. Wood and U. Wood-Sichra. Spatial Production Allocation Model (SPAM) 2000 Version 3 Release 1. Available at www.harvestchoice.org/production/systems/distribution/.

Watt, M.S., D.J Kriticos., S. Alcaraz, A. Brown and A. Leriche. "The Hosts and Potential Geographic Range of Dothistroma Needle Blight." *Forest Ecology and Management* 257 (2009): 1505-1519.

Woodward, F.I. *Climate and Plant Distribution*. Cambridge: Cambridge University Press, 1987.