

1 **A risk assessment framework for seed degeneration: Informing an**
2 **integrated seed health strategy for vegetatively-propagated crops**

3
4 S. Thomas-Sharma, J. Andrade-Piedra, M. Carvajal Yepes, J. F. Hernandez Nopsa, M. J. Jeger,
5 R. A. C. Jones, P. Kromann, J. P. Legg, J. Yuen, G. A. Forbes, K. A. Garrett

6
7 first, fourth, and eleventh authors: Department of Plant Pathology, Kansas State University, KS,
8 USA;

9 first author: Department of Plant Pathology, University of Wisconsin-Madison, WI, USA;

10 second author: International Potato Center, Lima, Peru;

11 third author: International Center for Tropical Agriculture, Cali, Colombia;

12 fourth and eleventh authors: Institute for Sustainable Food Systems, Plant Pathology

13 Department, and Emerging Pathogens Institute, University of Florida, Gainesville, FL, USA;

14 fifth author: Centre for Environmental Policy, Imperial College London, UK;

15 sixth author: Institute of Agriculture, University of Western Australia, Crawley, Australia;

16 seventh author: International Potato Center, Quito, Ecuador;

17 eighth author: International Institute of Tropical Agriculture, Dar es Salaam, Tanzania;

18 ninth author: Department of Forest Mycology and Plant Pathology, Swedish University of
19 Agricultural Sciences, Uppsala, Sweden;

20 tenth author: International Potato Center, Kunming, China.

21
22 Corresponding authors: K. A. Garrett and S. Thomas-Sharma

23 E-mail addresses: karengarrett@ufl.edu and thomasuga@gmail.com

24 **ABSTRACT**

25 Pathogen build-up in vegetative planting material, termed seed degeneration, is a major problem
26 in many low-income countries. When smallholder farmers use seed produced on-farm or
27 acquired outside certified programs, it is often infected. We introduce a risk assessment
28 framework for seed degeneration, evaluating the relative performance of individual and
29 combined components of an integrated seed health strategy. The frequency distribution of
30 management performance outcomes was evaluated for models incorporating biological and
31 environmental heterogeneity, with the following results. (1) On-farm seed selection can perform
32 as well as certified seed, if the rate of success in selecting healthy plants for seed production is
33 high; (2) When choosing among within-season management strategies, external inoculum can
34 determine the relative usefulness of ‘incidence-altering management’ (affecting the proportion of
35 diseased plants/seeds) and rate-altering management (affecting the rate of disease transmission in
36 the field); (3) Under severe disease scenarios, where it is difficult to implement management
37 components at high levels of effectiveness, combining management components can produce
38 synergistic benefits and keep seed degeneration below a threshold; (4) Combining management
39 components can also close the yield gap between average and worst-case scenarios. We also
40 illustrate the potential for expert elicitation to provide parameter estimates when data are
41 unavailable.

42

43 *Additional keywords:* banana, cassava, environmental heterogeneity, positive selection, potato,
44 root crops, seed degeneration, seed health, simulation models, sweetpotato, tuber crops,
45 vegetative propagation, yam.

46 INTRODUCTION

47 In vegetatively-propagated crops, pathogens tend to accumulate if planting material is drawn
48 from within a crop population over multiple generations, resulting in significant quality and yield
49 losses. This problem, termed seed degeneration (where ‘seed’ refers to vegetative planting
50 material), occurs commonly when certified, disease-free planting material is scarce and/or
51 expensive, as is the case in many low-income countries (Gibson and Kreuze 2014; Thomas-
52 Sharma et al. 2016) and for some specialty crops (Gergerich et al. 2015). An integrated seed
53 health strategy (Thomas-Sharma et al. 2016) is needed to address seed degeneration, drawing on
54 management components that are currently available to farmers, or can be made available in the
55 near future. We present a risk assessment framework for seed degeneration in vegetatively-
56 propagated crops, designed to support the development of strategies for integrating management
57 components.

58 Seed degeneration is affected by many biophysical factors such as the susceptibility of a
59 variety, the abundance of alternative hosts (e.g., weeds), the roles and efficiencies of vectors,
60 regional inoculum availability, and the conduciveness of weather for disease development and
61 spread. Processes such as reversion, where seed obtained from infected mother plants is healthy,
62 can reduce seed degeneration in sweetpotato (Gibson et al. 2014), potato (Bertschinger 1992),
63 and cassava (Fargette et al. 1996; Gibson and OtimNape 1997). The etiology of seed
64 degeneration is often specific to a crop and geographical region. Cassava mosaic geminiviruses
65 (CMGs) and Cassava brown streak viruses (CBSVs) are major causes of degeneration in East
66 Africa (Legg et al. 2015), while viruses associated with cassava frogskin disease are the main
67 causes of degeneration in South America (Carvajal-Yepes et al. 2014). For potato, viruses are a
68 major cause of seed degeneration around the world (Thomas-Sharma et al. 2016), while latent

69 tuber infections of the bacterial wilt pathogen, *Ralstonia solanacearum*, are a major problem in
70 tropical and subtropical countries (Mwangi et al. 2008), and the fungal pathogen *Rhizoctonia*
71 *solani* is a problem at high altitudes in the Andes (Fankhauser 2000).

72 The use of seed certified to be disease-free or to have high health status (hereafter
73 referred to as certified seed) is often recommended as the primary management strategy to
74 counter on-farm seed degeneration (Frost et al. 2013; Gergerich et al. 2015; Thomas-Sharma et
75 al. 2016). Examples of such formally regulated systems include the US National Clean Plant
76 Network (www.nationalcleanplantnetwork.org), which supplies seed material for many fruit
77 crops, the Wisconsin Seed Potato Improvement Association (www.potatoseed.org), which
78 supplies seed potato in Wisconsin, USA, and the Netherlands General Inspection Service for
79 Agricultural Seeds and Seed Potatoes (www.nak.nl), which certifies seed potatoes from the
80 Netherlands for global export. However, specialized programs designed to produce healthy seed
81 are rarely used by smallholder farmers in low-income countries (Thiele 1999; Thomas-Sharma
82 et al. 2016). In many low-income countries, 80-95% of seed is routinely obtained from informal
83 seed sources with poor or unknown seed health status (Mallowa et al. 2006; McGuire and
84 Sperling 2016; Thomas-Sharma et al. 2016). Thus, one focus for the application of this new risk
85 assessment framework is the context of low-income countries, especially food-security crops
86 such as banana and plantain, cassava, potato, sweetpotato, and yam.

87 Evaluating management components, and the potential synergies from combining them,
88 is a key part of a risk assessment framework for seed degeneration. Synergies can be evaluated in
89 terms of reductions in disease, or increases in yield, that are greater when management
90 components are combined compared to the sum of the individual component effects. Seed
91 degeneration can be managed by limiting epidemics in the field, using components such as:

92 certified seed; host resistance; roguing; selection of seed, cuttings or plants (referred to as seed
93 selection); and management of vectors, pathogens, and alternative hosts (Blomme et al. 2014;
94 Legg 1999; Thomas-Sharma et al. 2016). Grouping management components by their mode of
95 action can facilitate decision-making and generalization of results. One approach is to group
96 management components into those reducing initial inoculum for each new planting, those
97 reducing the rate of disease spread within a crop, and those reducing the time of exposure of the
98 crop (Berger 1977). Another approach is to group components based on their selectivity against
99 pathogens and whether they manage internal or external inoculum sources (Jones 2006). To
100 compare the performance of management components in this risk assessment framework for seed
101 degeneration, we group components as incidence- and rate-altering management components, a
102 logical distinction based on the structure of our models. Incidence-altering management
103 components affect the availability of inoculum from host material in the field or seed lot. For
104 example, roguing (removal of symptomatic plants from the field) affects disease incidence in the
105 field, while seed selection (selecting asymptomatic/least-symptomatic plants as the seed source
106 each season) and use of certified seed, alter disease incidence in the seed lot. Rate-altering
107 management strategies affect the rate of spread of disease in a field, and include strategies such
108 as use of host resistance and vector or pathogen management. (Management of alternative hosts
109 could be treated as reducing inoculum availability from hosts, or as reducing the rate of spread,
110 depending on the goals of an analysis.)

111 The importance of seed degeneration and the high cost of multi-year field experiments to
112 support empirical analyses have motivated several studies using analytical and simulation
113 models to better understand the process of seed degeneration. Several of these studies have
114 focused on management strategies for cassava diseases, and vector dynamics (Fargette and Vie

115 1995; Holt et al. 1997), illustrating how small management improvements can reduce the risk
116 that a cropping system approaches a threshold for rapid disease increase. In cassava mosaic
117 disease (CMD), weather conditions can affect symptom expression (Gibson and OtimNape
118 1997), so symptom-based management like roguing and seed selection can have variable success
119 rates from one season to the next. The level of farmer skill in identifying symptoms, especially
120 for pathogens that produce cryptic foliar symptoms, e.g., cassava brown streak disease (CBSD),
121 is another source of variation from one field to another. McQuaid et al. (2016) evaluated the
122 likely performance of roguing for CBSD in seed multiplication sites, showing the potential in
123 sites with low disease pressure. In a more general modeling study, van den Bosch et al. (2007)
124 found that management components such as *in vitro* propagation, high accuracy cutting
125 selection, and use of tolerant varieties, can inadvertently select for virus strains that build-up a
126 high titer in host plants. New analysis of how strategic integration of management components
127 enhances management performance can build on these studies.

128 Weather is another factor determining the rate of seed degeneration. Viral degeneration
129 of seed potato is lower at high altitudes (Rahman and Akanda 2008), due at least in part to lower
130 virus and vector activity (Fankhauser 2000), and higher rates of reversion or efficiency of
131 autoinfection (Bertschinger 1992). In a fine-scale forecasting model of potato viruses,
132 Bertschinger et al. (1995) used daily temperature measurements to determine host growth rates
133 and vector dynamics, predicting the number of infected progeny seed. In most other models of
134 seed degeneration, however, weather is implicitly addressed in vector dynamics, and weather
135 variability is rarely considered. Understanding the effect of season-to-season weather variability
136 is important for evaluating seed degeneration risk, and understanding climate change scenarios,

137 e.g., where increased population growth of potato virus vectors is predicted for summer crops in
138 parts of South Africa (van der Waals et al. 2013).

139 There are many potential goals for model development, such as providing a good
140 approximation to reality, precise predictions, or general insights into a phenomenon. Typically
141 models will compromise one of these objectives in the pursuit of others (Gross 2013; Levins
142 1966). Our goal in developing a risk assessment framework for seed degeneration is to provide a
143 *general* assessment of the performance of different management approaches, as well as a
144 framework that can be adapted to applications for specific pathosystems. We build on the
145 modeling studies of seed degeneration discussed above, with an emphasis on evaluating the
146 effects of both weather variability and variability in management implementation. Thus the
147 benefits of combining management components in an integrated seed health strategy can be
148 explored under different weather scenarios.

149 The limited data available related to the extent and variability of management component
150 adoption, especially in scenarios where seed degeneration is a problem (e.g., low-income
151 countries), can be a challenge for model parameterization. Studies generally report small-scale,
152 site-specific estimates, so there is little information to guide scaling up consideration to regional
153 or larger extents. In many applications where decisions have to be made despite severe data
154 limitations, such as conservation biology, the use of expert opinion to fill information gaps has
155 gained momentum (Mac Nally 2007; Martin et al. 2005; Yamada et al. 2003). ‘Expert
156 elicitation’ is the systematic collection of the wealth of information integrated into scientists’
157 opinions through the course of their studies of particular systems (Knol et al. 2010). Use in plant
158 pathology has generally been limited to applications such as the use of expert knowledge for
159 cluster sampling of disease incidence (Hughes and Madden 2002). We explored expert elicitation

160 as a tool to provide the frequency distribution of likely parameter values (such as the level of
161 disease resistance deployed) in India and Africa, along with information about the uncertainty
162 due to lack of knowledge about these systems. Because expert elicitation can provide
163 information about the deployment of a management component across farms in a region, the data
164 it provides can be used to scale up model results to evaluate regional management performance.

165 We develop here a general risk assessment framework for seed degeneration, designed to
166 inform an integrated seed health strategy for vegetatively-propagated crops (Thomas-Sharma et
167 al. 2016). The objectives of the study were to (1) build on current theoretical understanding of
168 seed degeneration by including stochasticity of both environmental factors and management
169 components, (2) evaluate scenarios where integrated seed health strategies would be more and
170 less successful, and (3) explore the use of expert elicitation as a method to complement
171 traditional empirical data. We used the framework to ask a set of key questions. (1) Certified
172 seed use is sometimes viewed as a “silver bullet” for managing seed degeneration, yet is
173 unavailable to many farmers. For what scenarios can on-farm management perform as well as
174 certified seed use? And for what scenarios is certified seed use of little value without on-farm
175 management? (2) Given the resource limitations of many farmers in low-income countries, is
176 there an epidemiological basis to choose among within-season management components? Which
177 management components would perform better in the presence or absence of external inoculum?
178 (3) Some methods such as seed selection may present challenges for achieving high levels of
179 effectiveness of implementation, due to cryptic symptoms or lack of farmer experience. Farmers
180 may also choose to plant a mixture of healthy and infected seed when healthy seed is limited and
181 reversion possible (Holt et al. 1997). Can combining management components reduce the
182 minimum effectiveness of implementation required for successful seed degeneration

183 management? (In this study, ‘effectiveness’ is generally used to refer to the effectiveness of
184 implementation of a management component, such as the degree of disease resistance, and
185 differentiated from the effect of management on yield, termed management ‘performance’.) (4)
186 In a development context, the focus may lie not only on the average performance of strategies,
187 but also on the tail of the distribution of performance, the farmers who may be experiencing least
188 benefit from strategies. A stochastic model allows evaluation of the relative performance of
189 strategies across the distribution. How should management be modified to reduce losses at the
190 tail?

191

192 **METHODOLOGY**

193 *Overview*

194 Purpose: Environmental conditions are important risk-factors for seed degeneration, affecting the
195 build-up of pathogens in a plant, the spread of disease within a field and across large geographic
196 areas, and the usefulness of management practices. A new focus in this framework on the
197 conduciveness of weather for seed degeneration, and variability in weather across cropping
198 seasons, is designed to complement insights from previous seed degeneration models. The
199 effectiveness with which management components are implemented by a farmer can also be
200 variable from one season to the next, depending on the available resources. We address seed
201 degeneration defined as pathogen build-up in seed material, while other physiological factors
202 (such as seed age and physical damage/abnormalities) which can also lower the quality of seed
203 are not considered (Thomas-Sharma et al. 2016). This risk assessment framework for seed
204 degeneration is designed to be broadly applicable to vegetatively-propagated crops/pathosystems
205 and to capture the key seasonal dynamics of seed degeneration (Fig. 1). While this is not an

206 agent-based model, we generally followed a model description format recommended for agent-
207 based models (Grimm et al. 2010), to enhance clarity and reproducibility. An interactive
208 interface, built by Y. Xing and S. Thomas-Sharma using the Shiny package in R, allows users to
209 experiment directly with the models described here, by accessing the code used in this analysis.
210 It is available at <https://yanru-xing.shinyapps.io/SDAppvX1/>.

211 Scales and state variables: The model time-step is a season (s), defined as a ‘vegetative
212 generation’, i.e., the time between planting and seed collection during which management
213 decisions are made. For crop species where seeds are collected on a different time scale than
214 harvest of the food crop (e.g., banana or sweetpotato), the production of seed (e.g., banana
215 suckers and sweetpotato vines) and the production of food (e.g., fruit and storage roots) can be
216 considered separately. Seed degeneration is modeled in an individual field without spatially
217 explicit structure, over multiple seasons. Plant and seed populations are characterized by the
218 number or proportion of healthy and diseased individuals, determining the resulting yield loss
219 each season. The state variables are healthy (HP_s) and diseased (DP_s) plant numbers, healthy
220 (HS_s) and diseased (DS_s) seed proportions, end of season yield (Y_s) and end of season percent
221 yield loss (YL_s) (Table 1).

222 Process overview and scheduling: The model includes five processes that occur every
223 season: host infection, host removal, seed formation, seed selection, and seed choice (Fig. 1).
224 The effects of the following management strategies are evaluated: use of certified seed, host
225 resistance, roguing, seed selection, and vector or pathogen management. (Management of
226 alternative hosts could be evaluated explicitly by incorporating an additional model component,
227 or implicitly as part of the effects of vector or pathogen management.)

- 228 (1) Host infection, or disease transmission, increases disease incidence in the field, and is a
229 function of the disease incidence in the seed and the availability of external inoculum. The
230 rate of disease transmission is determined by the maximum seasonal disease transmission,
231 the degree to which weather conditions are disease-conducive, any external inoculum
232 present, and the levels of rate-altering management applied in the field (i.e., host resistance,
233 vector or pathogen management). A subset of analyses highlight the greater impact of early-
234 season infections compared to late-season infections. Good proxies for the level of external
235 inoculum are challenging to obtain; in this framework, we included external inoculum as a
236 factor that acts comparably to the presence of infected plants within the field.
- 237 (2) Host removal occurs by roguing, where diseased plants are removed from the population
238 (death due to disease is treated as minimal). In a subset of analyses, where early- and late-
239 season infections are considered, we also highlight the effects of roguing conducted early
240 versus late in the season. Specifying a minimum yield (*minY*) greater than zero supports
241 analysis of the yield penalty due to roguing (when diseased plants produce usable yield). Any
242 compensatory yield effects when roguing is applied (when surrounding plants compensate
243 for yield loss; Salazar 1996) have not been considered.
- 244 (3) We use the term ‘seed formation’ to describe the production of seed, where the health of the
245 mother plant determines the health of the seed. During seed formation, reversion causes a
246 proportion of diseased plants to become disease-free, producing healthy seed. Diseased
247 plants may produce less seed, contributing less diseased seed to the total on-farm seed
248 produced.
- 249 (4) Seed selection is represented by a change in the proportion diseased seed produced as a result
250 of selecting against diseased plants as the seed source each season. We do not explicitly

251 describe a distinction often made in seed selection, between positive selection (selection of
252 asymptomatic plants for seed under high disease intensity) and negative selection (rejection
253 of symptomatic plants for seed under low disease intensity). In this model, the proportion of
254 diseased seed is reduced, which might be due to either positive or negative selection.

255 (5) Seed choice affects the proportion of on-farm seed that is combined with certified seed and
256 used in the next season. The model is not spatially explicit, so the degree of mixing among
257 all seed planted in the field is assumed not to have an important effect.

258 *Design concepts*

259 Stochasticity: Two general components are stochastic in this model: seasonal weather-
260 conduciveness for disease, and the effectiveness with which management strategies (vector or
261 pathogen management, seed selection, and roguing) are implemented. The parameters describing
262 these components are the weather index (W), the proportional change in infection rate due to
263 vector or pathogen management (M), the proportional selection against diseased seed (Z), and
264 the proportion diseased plants remaining after roguing (A). Each of these follows a normal
265 distribution truncated between 0 and 1, where realizations below 0 are treated as 0, and
266 realizations above 1 are treated as 1.

267 For the weather index, the mean represents mean climatic conditions, and variation
268 around the mean represents season-to-season variability in conduciveness to disease. For vector
269 or pathogen (or alternative host) management, the mean represents mean effectiveness with
270 which practices are applied and stochasticity captures season-to-season variability, due to timing
271 of application or incomplete control (e.g., changes in the timing and choice of insecticides for
272 vector management). For management practices based on symptom recognition (roguing and
273 seed selection), the mean indicates the mean effectiveness with which the practice is applied

274 during a season. Stochasticity captures both variability in symptom expression (e.g., due to
275 variability in timing of infection among seasons, or delayed symptom development) and
276 variability in farmers' skill in recognizing symptomatic plants. Thus stochasticity in these
277 analyses generally represents what Oberkampf et al. (2004) refers to as 'aleatory uncertainty'.
278 Similar analyses could also be interpreted in terms of 'epistemic uncertainty' or uncertainty due
279 to lack of system knowledge (Oberkampf et al. 2004), or a combination of these two types of
280 uncertainty.

281 Calibration and rate of disease transmission: We conceptualize β as the maximum rate of
282 disease transmission during the growing season, associated with a scenario where there are no
283 limiting factors for disease spread (i.e., when there is no vector or pathogen management, a
284 highly susceptible host is planted, and the weather is highly disease-conducive). This rate is not
285 necessarily intuitive, because it is multiplied by the number of diseased and healthy plants, in
286 addition to being modified by parameters reflecting the effects of vector or pathogen
287 management, host resistance, and weather. β is determined by vector and pathogen attributes and
288 other dispersal characteristics, and is interpreted in the context of this general framework as
289 reflecting the maximum rate in the absence of limiting factors. In most simulation experiments,
290 we took $\beta = 0.02$ as the maximum disease transmission rate per season. After exploring the
291 behavior of β at high and low starting levels of infection with and without management
292 strategies, $\beta = 0.02$ was selected to provide a range of outcomes for evaluation. Substantially
293 lower or higher values of β resulted in consistent lack of disease, or immediately high disease
294 levels, respectively. Identifying a value of β through this type of calibration met the needs of our
295 general analysis. However, when developing a more precise application of this framework for

296 managing a specific crop, calibrating β for the pathosystem and relevant environments will be an
297 important step.

298 Observations: Model output for a given parameter combination includes summary
299 statistics (mean, 5th percentile, and 95th percentile) for the state variables, and timing for renewal
300 with certified seed, i.e., the first time point at which the proportion of healthy seed falls below a
301 threshold value, where a threshold proportion of 0.7 was used in examples.

302 *Details*

303 Initialization: The initial proportion of healthy seed (pHS_0) determines the starting infection level
304 in the field, in the first time-step. A relatively low proportion initial infection (0.2) was assumed
305 in most scenarios, and in some cases was compared with a high proportion infection (0.8). Such
306 high pathogen incidence in planting material is a common scenario in low-income countries
307 where farmers routinely use seed of poor health status (Gildemacher et al. 2009), or fields have
308 high disease incidence, making it difficult to select disease-free planting material (Legg 1999).

309 Input data: The current application of the model does not depend on external weather
310 data. However, for more specific applications, the weather index parameter could be defined as a
311 function of a set of observed weather variables relevant to a particular pathosystem.

312 Submodels: There are four submodels that incorporate the effects of weather and
313 management on the state variables (details in Supplementary material S1). The first submodel
314 determines the number of healthy plants at the end of a season as a function of the proportion of
315 diseased seed at the beginning of the season. The second submodel determines the number of
316 healthy seed produced in that season. The third determines yield (in terms of food production) as
317 a function of the proportion of diseased plants. The fourth determines the proportion healthy seed
318 used in the next season.

319 ***Simulation experiments***

320 Simulation experiments were implemented in the R programming environment (R Core Team
321 2015). Experiments were designed to evaluate the effect of environment, and individual or
322 combined management strategies on yield loss due to seed degeneration, to address the questions
323 posed in the introduction. Each parameter combination was evaluated in 2000 simulations. The
324 maximum (*maxY*) and minimum (*minY*) attainable yields were set at 100 and 0 units,
325 respectively. Some parameters were evaluated at contrasting levels, while other parameters were
326 set to the default values in Table 2 when their effects were not being evaluated. For example,
327 management practices were set to minimum values (i.e., 0), without stochasticity, unless the
328 impact of different levels of effectiveness of implementation was being evaluated. The default
329 values of parameters were selected such that contrasting outcomes could be evaluated. All results
330 are represented such that 0 indicates lack of a management component and 1 indicates complete
331 effectiveness of implementation. (The effects of roguing, seed selection, vector management, and
332 host resistance are likewise described in the results in terms of the effectiveness of
333 implementation, rather than in terms of their corresponding parameter definitions in the model
334 Table 2). The standard deviation for stochastic variables was set to 0.3 and 0.1 for high and low
335 variability scenarios, respectively. Short- (5 season) and long-term (10 season) effects on yield
336 loss were studied.

337 ***Parameterization based on expert elicitation***

338 The risk assessment framework described to this point is designed to evaluate risk at a particular
339 field, given the environment and management decisions implemented. Expert elicitation was
340 used to assess the adoption rates for individual management components by farmers in a region,
341 as a first step toward scaling up individual farm risk assessments. In total, twenty-five experts

342 (across crops and geographical regions) provided estimates of the frequency with which different
343 management components were implemented with a particular level of effectiveness. For
344 example, experts estimated the field acreage in each of 10 disease resistance categories in
345 regions of Africa and India. The seed degeneration model described above was used to evaluate
346 outcomes for an individual field, providing the frequency of potential outcomes for a given
347 scenario defined by a set of parameter values. To supplement individual field evaluation, the
348 expert elicitation data provide estimates of the frequency with which different scenarios occur.
349 The data from expert elicitation were used to *partially* calibrate the frequency distribution of
350 yield loss in the risk assessment framework for seed degeneration. Expert elicitation provided
351 relatively high confidence information about the frequency with which farmers used particular
352 management techniques, but did not provide high confidence estimates related to underlying
353 transmission rates (because of the inherent difficulty in estimating transmission rates from
354 personal observations). Thus, expert elicitation made this general analysis relatively more
355 realistic, by indicating how likely different scenarios were to occur, but did not provide a precise
356 estimate of yield outcomes. The details of the methods employed in expert elicitation are in
357 Supplementary material S2.

358

359 **RESULTS**

360 *Effect of weather on long-term yield loss*

361 The effect of disease-conducive weather conditions on long-term yield loss was first illustrated
362 in the absence of management, and external inoculum, with other parameters set to default. As
363 expected, highly disease-conducive weather causes yield loss to rise quickly, while, under
364 marginally disease-conducive weather, it rises relatively more slowly and has the potential to

365 stay at an acceptable level (Fig. 2). Season-to-season variability in weather causes seasonal
366 fluctuations in yield loss. Under marginally disease-conducive weather, this variability can cause
367 long-term yield reductions to be very high and comparable to that in highly disease-conducive
368 weather conditions.

369 *Effect of individual management practices on yield loss*

370 The effect of individual management practices on short-term yield loss varies with the degree to
371 which weather is disease-conducive (Fig. 3). As disease conduciveness increases, management
372 practices provide less reliable yield loss reduction. For all cases illustrated, under highly disease-
373 conducive conditions, yield loss reaches nearly 100% when the proportional effectiveness of
374 implementation of management practices is low (0-0.2). The effects of the incidence-altering
375 management practices such as roguing, seed selection, and certified seed use are similar to each
376 other. As expected based on the model structure, rate-altering management strategies, such as
377 vector or pathogen management and host resistance, had the same outcome for a given
378 effectiveness of implementation (*not shown separately*).

379 In the absence of external inoculum, strategies such as roguing, use of certified seed and
380 seed selection could substantially reduce yield loss when implemented at 0.2-0.4 proportional
381 effectiveness, under marginally disease-conducive conditions (Fig. 3). Rate-altering management
382 strategies, however, required higher levels of proportional effectiveness of implementation (0.4-
383 0.6) to provide a comparable effect on yield loss. Even when rate-altering management strategies
384 were implemented at ‘complete’ proportional effectiveness (i.e., at 1), in marginally disease-
385 conducive weather conditions, a low level of yield loss (~10%) was observed (Fig. 3A). This was
386 because it took more than 5 seasons for rate-altering management to reduce yield loss levels to
387 zero (*data not shown*). Depending on weather conduciveness and resistance levels, management

388 practices such as roguing, use of certified seed and seed selection were thus 20-40% more
389 beneficial than rate-altering management strategies, in the absence of external inoculum (Table
390 3).

391 When external inoculum is present, however, incidence-altering management was less
392 successful than rate-altering management strategies, reversing the ranking observed in the
393 absence of external inoculum (Fig. 4A, B). When both seed selection and vector or pathogen
394 management were implemented at 0.6 proportional effectiveness, the use of vector or pathogen
395 management in the presence of external inoculum (Fig. 4D) resulted in a relatively slower
396 increase in long-term yield loss compared to seed selection (Fig 4C).

397 *Effect of combining management strategies on yield loss*

398 The minimum level of effectiveness of implementation for a management component to keep
399 long-term yield loss below 10% (in the absence of external inoculum), changed with the level of
400 resistance used (Table 3). Under highly disease-conducive weather conditions, when susceptible
401 varieties were grown, vector or pathogen management, roguing, seed selection and external
402 certified seed had to be used at 0.9, 0.7, 0.7 and 0.6 proportional effectiveness, respectively, to
403 maintain yield loss <10%. If a resistant variety was used, however, this minimum effectiveness
404 of implementation could be lowered (Table 3). In scenarios where starting infection is high and
405 weather is highly conducive for disease, seed selection is insufficient to keep yield loss below
406 10% in susceptible varieties (*data not shown*).

407 Combining management strategies is also useful to delay the need for seed renewal from
408 off-farm certified sources (Table 4). Consider a scenario where renewing seed material with off-
409 farm certified seed becomes necessary when the healthy seed proportion falls below a threshold
410 of 0.7 (which corresponds to approximately 30-40% yield loss depending on conduciveness of

411 weather). In the presence of external inoculum and highly disease-conducive weather conditions,
412 seed renewal was necessary every season when seed selection and vector or pathogen
413 management were practiced individually, but when these practices were combined seed renewal
414 was not necessary for ~12 seasons. In this case, there was strong synergy in the sense that the
415 time to seed renewal for the combined management was substantially larger than the sum of the
416 times to renewal for the two components individually.

417 *Effect of season-to-season variability in weather and management practices*

418 Under high proportional effectiveness of implementation (>0.8), high season-to-season
419 variability in vector or pathogen management (*data not shown*) or seed selection resulted in
420 greater yield loss under highly disease conducive weather conditions (Fig. 5). In marginally
421 disease-conducive weather (<0.2) and low proportional effectiveness of implementing seed
422 selection (<0.2), high variability in selection (Fig. 5 B, D) resulted in lower yield loss than low
423 variability scenarios (Fig. 5 C, D). This was because, given the model structure, at low
424 effectiveness of implementation, variability resulted in a higher proportion of healthy plants
425 being incorporated. Conversely, under high effectiveness of implementation, variability in
426 selection resulted in the incorporation of more diseased plants. These trends were more
427 predominant when the starting infection-levels were high (Fig. 5 C, D).

428 Season-to-season variability in weather and management practices resulted in variable
429 levels of yield loss (Table 5). In addition to the mean outcomes, we considered the near worst-
430 case outcomes (5th percentile) and the near best-case outcomes (95th percentile). In the near
431 best-case outcome, by implementing seed selection at 0.6 proportional effectiveness for a variety
432 with resistance at level 0.6 out of 1.0, a farmer incurred a yield loss of 16% under highly disease-
433 conducive weather conditions, in the presence of external inoculum (Table 5). However, in the

434 worst-case outcome, implementing management components at the same level of effectiveness
435 resulted in 50% yield loss (Table 5). In the absence of external inoculum, combining seed
436 selection and host resistance resulted in <5% yield loss in best-, worst-case and average
437 outcomes (Table 5).

438 *Use of expert elicitation to provide input for crop-specific analyses*

439 In the absence of information about geographic deployment of resistance in cassava, each level
440 of resistance might be considered equally likely, as in an uninformative prior in Bayesian
441 analysis. For a uniform distribution of resistance deployment, model predictions for yield loss in
442 a region would be considerably lower than are likely to be observed, given the rarity of
443 resistance deployment reported in expert elicitation. Crop-specific acreage information obtained
444 from experts (Fig. 6A) can be used to estimate regional yield loss. The resulting modified yield
445 loss distribution (Fig. 6C) is one step more realistic for cassava in Africa and India, in this
446 illustration for marginally disease-conducive weather scenarios.

447

448 **DISCUSSION**

449 The seed degeneration risk assessment framework was designed to identify scenarios in low-
450 income countries where on-farm management components may be useful, and where they may
451 be absolutely necessary to slow or reverse seed degeneration. We observed that:

452 *(1) On-farm seed selection can perform as well as certified seed use, if the rate of success in*
453 *selecting healthy plants is high.* Using the risk assessment framework for seed degeneration, we
454 illustrate how roguing and seed selection can perform as well as use of certified seed (Fig. 3).
455 For many pathosystems, achieving a suitably high rate of success in symptom-recognition is
456 challenging when symptoms are cryptic or variable. If the effectiveness of implementation is

457 low, high yield loss may result despite practicing seed selection (Fig 5D). For cassava seed
458 degeneration caused by CBSD, above-ground symptoms are cryptic (Legg et al. 2011), while for
459 CMD, symptom expression is much more reliable but depends partially on seasonal weather
460 conditions (Gibson and OtimNape 1997). The usefulness of seed selection has been
461 demonstrated for some diseases. For viral diseases and bacterial wilt in potato, farmer-managed
462 trials of seed selection resulted in a ~30% yield increase, with lower disease incidence
463 (Gildemacher et al. 2011; Schulte-Geldermann et al. 2012). Seed selection also increased the
464 tuberous root yields of CMD-susceptible cassava varieties (Mallowa et al. 2006). In parts of
465 western Kenya where CMD is in its post epidemic phase, there is a resurgence of local landraces
466 that are CMD-susceptible, partly because farmers choose the most vigorous plants as seed
467 sources (Mallowa et al. 2006). Farmers may decide against roguing when diseased plants
468 produce usable yield, limiting the practical usefulness of roguing (Legg et al. 2015; Mallowa et
469 al. 2011; Sisterson and Stenger 2013). Under such situations, restricting roguing to early in the
470 season (Supplementary material S3) and coordinating roguing over regional scales (Sisterson and
471 Stenger 2013) can increase the benefits and potentially the incentives for roguing. Additionally,
472 the model treats certified seed material as completely disease-free. Deep-sequencing techniques
473 have revealed that many plant viruses are yet to be described (Kreuze et al. 2009). Thus accurate
474 certification depends on the characterization of a crop virome which can evolve over time and is
475 currently unstudied for crops in many geographical regions. Finally, it is important to remember
476 that, in many low-income countries, farmers have limited or no access to certified seed, and seed
477 selection with even sub-optimum efficiency may provide yield benefits (Holt et al. 1997).

478 *(2) When choosing among within-season management components, external inoculum*
479 *can determine the relative usefulness of incidence- and rate-altering management. Management*

480 practices such as vector or pathogen management and host resistance can be particularly useful
481 when external inoculum is high and plants are potentially at risk of rapid infection from
482 surrounding fields (Fig. 4). In sweetpotato, the proximity of and level of inoculum in
483 surrounding fields (a function of the level of host resistance in surrounding fields), affected the
484 incidence of sweetpotato viral disease (Aritua et al. 1999). Fargette and Vie (1995) suggested
485 that phytosanitation (by selection of cuttings and roguing) would be more useful in areas with
486 high inoculum levels or susceptible cultivars, because under low inoculum pressure, the use of
487 resistant cultivars with reversion would sufficiently manage the disease. However, where local
488 inoculum levels are very high, plants would get infected quickly, making selection of cuttings
489 and roguing less feasible. In our evaluation, under low external inoculum (e.g., when crop fields
490 are isolated from each other), and when 10% of the cuttings underwent reversion, incidence-
491 altering management was better for managing degeneration than rate-altering management
492 strategies. In post CMD epidemic areas (where there is reduced inoculum), seed selection along
493 with reversion and natural cross protection by mild viruses together allow farmers to cultivate
494 locally preferred, CMD-susceptible varieties (Mallowa et al. 2006).

495 *(3) For severe disease scenarios, when implementing management components at high*
496 *levels of effectiveness is difficult, combining management components can produce synergistic*
497 *benefits and keep seed degeneration below a threshold. For Potato leafroll virus (PLRV)*
498 *management, the threshold at which vector management becomes necessary may be modified by*
499 *using resistant varieties (Difonzo et al. 1995). Our study illustrates how the time until renewal*
500 *with certified seed was needed was prolonged when seed selection and host resistance were*
501 *applied together (Table 4). In the presence of external inoculum, it is particularly clear that the*
502 *performance of the combined strategy would be greater than the combination of simple additive*

503 effects of individual components, demonstrating potential synergy. We make a simplifying
504 assumption that the time until renewal with certified seed, and the choice of how to integrate
505 management components, depends solely on yield loss. In reality, many socioeconomic factors
506 such as cost and incentives for management, stakeholder preferences, etc., should also be
507 considered to better understand the factors affecting renewal with certified seed, and adoption
508 rates of integrated management practices more broadly in low-income countries (Parsa et al.
509 2014).

510 *(4) Combining management components can close the yield gap between average and*
511 *near worst-case outcomes caused by weather and management heterogeneity.* In the context of
512 development, there may be particular concern for the worst-case outcomes, such as when
513 particularly disease-conducive years may drive vulnerable farmers out of business. High seasonal
514 fluctuations in weather in a geographic region can result in very high yield reductions in the
515 long-term, despite the region being categorized as marginally disease-conducive (Fig 2). The
516 increased frequency of extreme weather is predicted in many climate change scenarios and can
517 lower the performance of disease management practices (Garrett et al. 2013; Jones 2016;
518 Lamichhane et al. 2015). McQuaid et al. (2016) report scenario analyses where, to keep CBSD
519 infection below 10% in seed farms located in areas with lower inoculum pressure, roguing
520 needed to be conducted frequently (weekly or fortnightly intervals) and at a relatively high
521 success rate (70% or higher). Roguing susceptible varieties can however significantly reduce
522 yields compared to a ‘do nothing’ strategy, largely due to the elimination of plant populations
523 (Mallowa et al. 2011). In such cases where farmers are hesitant to remove plants or have
524 difficulty recognizing cryptic symptoms (Legg et al. 2011), infection rates may become too high
525 for roguing or seed selection to be successful. Yield losses can be much higher for near-worst

526 case outcomes than average outcomes, but can be improved by combining management practices
527 (Table 5).

528 This risk assessment framework for seed degeneration was designed to capture season-to-
529 season dynamics of degeneration, so the model aggregates within-season characteristics of
530 degeneration in most analyses. The model defines degeneration as an increase in incidence of
531 infected seed in the seed lot and does not capture an increase in pathogen load (e.g., virus titer)
532 within a plant. Other specific attributes of vectors such as inoculation efficiency, acquisition
533 efficiency, vector birth/death rates, etc., that are relevant at finer temporal resolution, have been
534 aggregated in the maximum seasonal transmission rate (β) and the parameters that modify it.
535 Thus, when calibrating the model for application to a particular crop species in a particular
536 location, developing good estimates of β and modifying parameters will be a key step.
537 Calculating the rate of disease transmission when there are no constraints for disease
538 development provides an estimate for β (e.g., scenarios where a highly susceptible variety is
539 grown under highly-conducive weather conditions and without management).

540 Another goal of this risk assessment framework for seed degeneration is flexibility in
541 adapting to the study of different pathogens causing seed degeneration. Although viruses are a
542 major cause of seed degeneration in many crops, in West African yam, the nematodes
543 *Scutellonema bradys*, *Pratylenchus coffeae*, and *Meloidogyne* spp. are major causes of
544 degeneration (Coyne et al. 2010). In bananas grown in Africa, many nematodes (e.g.,
545 *Radopholus similis* and *Pratylenchus coffeae*) and bacterial pathogens (e.g., *Xanthomonas*
546 *campestris* pv. *musacearum*) readily accumulate and spread via planting material (Blomme et al.
547 2014). In general, β may be substantially higher for vector-borne viruses compared to soil-borne
548 fungal pathogens. Also, the weather index directly modifies β and may be conceptualized as

549 directly affecting nematodes, fungal pathogens, or the dynamics of virus vectors. The potential
550 effectiveness of implementation of management components may vary widely for management
551 of vectors, fungi, bacteria, and nematodes, and can be modified accordingly.

552 We used expert elicitation to obtain parameter estimates for use in the seed degeneration
553 risk assessment model. Although expert elicitation cannot replace empirical experimentation, we
554 were interested in exploring expert elicitation as a tool to characterize the frequency of different
555 cropping scenarios in a region, which can then be updated as more direct observations become
556 available. A limitation of data from expert elicitation is its subjective nature, potentially
557 influenced by biases. Using data from expert elicitation, we were able to evaluate the relative
558 effects of other management components, taking into account expert estimates of how commonly
559 and with what level of effectiveness the management components were implemented.

560 This seed degeneration risk assessment framework was designed to answer general
561 questions about the relative performance of management components, alone and in combination,
562 and to provide a platform to answer ‘what-if’ questions for specific scenarios of a crop,
563 pathosystem, and geographic region. For any given pathosystem, implementing the framework
564 can also help to identify key gaps in current knowledge, where parameter estimates are difficult
565 to obtain, that could be the focus of future field studies (Restif et al. 2012). For example, the
566 regional conduciveness of weather to disease could be evaluated based on general observations
567 of regional disease severity, keeping in mind that crop host availability can also be a limiting
568 factor for disease. If good models of weather effects on vector or pathogen dynamics are
569 available, these could be used to evaluate disease-conduciveness in a more flexible way, with
570 more potential to study the effects of weather variability and climate change, and to partition the
571 effects of weather and host abundance. As parameter estimates for a particular pathosystem

572 become available from field studies, the framework can be used to answer questions about the
573 time until renewal with certified seed becomes necessary, and how effectively management
574 components have to be implemented to keep yield loss below a threshold. Ongoing work with
575 the framework is aimed at expanding it to a regional scale in addition to analysis of individual
576 fields, through added information from the literature, new field studies, and expert elicitation.
577 Individual growers act in networks, within which they intentionally and unintentionally exchange
578 information, seed, and pathogens (Garrett 2012; Moslonka-Lefebvre et al. 2011; Shaw and
579 Pautasso 2014). Evaluating the structure of regional networks for the movement of seed may
580 help in targeting where extension and mitigation are most important (Hernandez Nopsa et al.
581 2015), and may also help to address the challenge of understanding the role of external inoculum
582 in disease risk within a field. Network concepts may also be extended to consider how landscape
583 structures create links among host species (Cox et al. 2013), an important factor for generalist
584 pathogens and vectors. Another important outcome of a regional framework would be regional or
585 larger-extent maps of the likely performance of different seed degeneration management
586 strategies, as an extension of concepts in species distribution mapping (Franklin 2009). Maps of
587 likely management performance can help to inform prioritization by policy makers and extension
588 groups.

589

590 **ACKNOWLEDGMENTS**

591 KAG, STS, and GAF developed the framework and model with input from all authors. STS and
592 KAG wrote the R code. STS performed the computational experiments. STS interviewed
593 experts for the expert elicitation component. STS, KAG, and GAF wrote the manuscript with
594 input from all authors. This work was supported by the CGIAR Research Program on Roots,

595 Tubers, and Bananas, the CGIAR Research Program on Climate Change, Agriculture and Food
596 Security, the Kansas Agricultural Experiment Station (Contribution no. ____), a NSF-BMGF
597 BREAD Idea Challenge Award to STS, US NSF Grant EF-0525712 as part of the joint NSF-NIH
598 Ecology of Infectious Disease program, US NSF Grant DEB-0516046, the National Institute for
599 Mathematical and Biological Synthesis (NIMBioS), and the University of Florida. We thank
600 scientists at the following institutions for their participation in the expert elicitation study:
601 Central Tuber Crop Research Institute, Thiruvananthapuram, India; Banana Research Station,
602 Kannara, India; Kerala Agricultural University (KAU), Thrissur, India; KAU,
603 Thiruvananthapuram, India; National Research Center for Banana, Tiruchirapalli, India;
604 International Institute of Tropical Agriculture (IITA), Dar es Salam, Tanzania; IITA Ibadan,
605 Nigeria; University of Queensland, Brisbane, Australia; International Potato Center (CIP), Quito,
606 Ecuador; CIP, Lima, Peru; CIP, Beijing, China; and Indian Agricultural Research Institute, New
607 Delhi, India.

608

609 LITERATURE CITED

- 610 Aritua, V., Legg, J. P., Smit, N., and Gibson, R. W. 1999. Effect of local inoculum on the spread
611 of sweet potato virus disease: limited infection of susceptible cultivars following
612 widespread cultivation of a resistant sweet potato cultivar. *Plant Pathology* 48:655-661.
- 613 Berger, R. D. 1977. Application of epidemiological principles to achieve plant disease control.
614 *Annual Review of Phytopathology* 15:165-183.
- 615 Bertschinger, L. 1992. Modelling of potato virus pathosystems by means of quantitative
616 epidemiology. PhD thesis. Swiss Federal Institute of Technology, Zurich, Switzerland.
- 617 Bertschinger, L., Keller, E. R., and Gessler, C. 1995. Characterization of the virus x temperature
618 interaction in secondarily infected potato plants using EPIVIT. *Phytopathology* 85:815-
619 819.
- 620 Blomme, G., Jacobsen, K., Ocimati, W., Beed, F., Ntamwira, J., Sivirihauma, C., Ssekiwoko, F.,
621 Nakato, V., Kubiriba, J., Tripathi, L., Tinzaara, W., Mbolela, F., Lutete, L., and
622 Karamura, E. 2014. Fine-tuning banana *Xanthomonas* wilt control options over the past
623 decade in East and Central Africa. *European Journal of Plant Pathology* 139:265-281.

- 624 Carvajal-Yepes, M., Olaya, C., Lozano, I., Cuervo, M., Castano, M., and Cuellar, W. J. 2014.
625 Unraveling complex viral infections in cassava (*Manihot esculenta* Crantz) from
626 Colombia. *Virus Research* 186:76-86.
- 627 Cox, C. M., Bockus, W. W., Holt, R. D., Fang, L., and Garrett, K. A. 2013. Spatial
628 connectedness of plant species: potential links for apparent competition via plant
629 diseases. *Plant Pathology* 62:1195-1204.
- 630 Coyne, D. L., Claudius-Cole, A. O., Kenyon, L., and Baimey, H. 2010. Differential effect of hot
631 water treatment on whole tubers versus cut sets of yam (*Dioscorea* spp.). *Pest*
632 *Management Science* 66:385-389.
- 633 Difonzo, C. D., Ragsdale, D. W., and Radcliffe, E. B. 1995. Potato leafroll virus spread in
634 differentially resistant potato cultivars under varying aphid densities. *American Potato*
635 *Journal* 72:119-132.
- 636 Fankhauser, C. 2000. Seed-transmitted diseases as constraints for potato production in the
637 tropical highlands of Ecuador. PhD thesis. Swiss Federal Institute of Technology, Zurich,
638 Switzerland.
- 639 Fargette, D., and Vie, K. 1995. Simulation of the effects of host resistance, reversion, and cutting
640 selection on incidence of African cassava mosaic virus and yield losses in cassava.
641 *Phytopathology* 85:370-375.
- 642 Fargette, D., Colon, L. T., Bouveau, R., and Fauquet, C. 1996. Components of resistance of
643 cassava to African cassava mosaic virus. *European Journal of Plant Pathology* 102:645-
644 654.
- 645 Franklin, J. 2009. Mapping species distributions: spatial inference and prediction. Cambridge
646 University Press, Cambridge, U.K.
- 647 Frost, K. E., Groves, R. L., and Charkowski, A. O. 2013. Integrated control of potato pathogens
648 through seed potato certification and provision of clean seed potatoes. *Plant Disease*
649 97:1268-1280.
- 650 Garrett, K. A. 2012. Information networks for disease: commonalities in human management
651 networks and within-host signalling networks. *European Journal of Plant Pathology*
652 133:75-88.
- 653 Garrett, K. A., Dobson, A. D. M., Kroschel, J., Natarajan, B., Orlandini, S., Tonnang, H. E. Z.,
654 and Valdivia, C. 2013. The effects of climate variability and the color of weather time
655 series on agricultural diseases and pests, and decision-making for their management.
656 *Agricultural and Forest Meteorology* 170:216-227.
- 657 Gergerich, R. C., Welliver, R. A., Gettys, S., Osterbauer, N. K., Kamenidou, S., Martin, R. R.,
658 Golino, D. A., Eastwell, K., Fuchs, M., Vidalakis, G., and Tzanetakis, I. E. 2015.
659 Safeguarding fruit crops in the age of agricultural globalization. *Plant Disease* 99:176-
660 187.
- 661 Gibson, R. W., and OtimNape, G. W. 1997. Factors determining recovery and reversion in
662 mosaic-affected African cassava mosaic virus resistant cassava. *Annals of Applied*
663 *Biology* 131:259-271.
- 664 Gibson, R. W., and Kreuze, J. F. 2014. Degeneration in sweetpotato due to viruses, virus-cleaned
665 planting material and reversion: a review. *Plant Pathology* 64:1-15.
- 666 Gibson, R. W., Wasswa, P., and Tufan, H. A. 2014. The ability of cultivars of sweetpotato in
667 East Africa to 'revert' from *Sweet potato feathery mottle virus* infection. *Virus Research*
668 186:130-134.

- 669 Gildemacher, P. R., Schulte-Geldermann, E., Borus, D., Demo, P., Kinyae, P., Mundia, P., and
670 Struik, P. C. 2011. Seed potato quality improvement through positive selection by
671 smallholder farmers in Kenya. *Potato Research* 54:253-266.
- 672 Gildemacher, P. R., Demo, P., Barker, I., Kaguongo, W., Woldegiorgis, G., Wagoire, W. W.,
673 Wakahiu, M., Leeuwis, C., and Struik, P. C. 2009. A description of seed potato systems
674 in Kenya, Uganda and Ethiopia. *American Journal of Potato Research* 86:373-382.
- 675 Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., and Railsback, S.F. 2010. The
676 ODD protocol: a review and first update. *Ecological Modelling* 221:2760-2768.
- 677 Gross, L. J. 2013. Use of computer systems and models. Pages 213-220 in: Levin, S.A., ed.
678 *Encyclopedia of Biodiversity, Second Edition, Volume 2*, pp. 213-220. Academic Press,
679 Waltham, MA.
- 680 Hernandez Nopsa, J. F., Daglish, G. J., Hagstrum, D. W., Leslie, J. F., Phillips, T. W., Scoglio,
681 C., Thomas-Sharma, S., Walter, G. H., and Garrett, K. A. 2015. Ecological networks in
682 stored grain: Key postharvest nodes for emerging pests, pathogens, and mycotoxins.
683 *BioScience* 65:985-1002.
- 684 Holt, J., Jeger, M. J., Thresh, J. M., and Otim-Nape, G. W. 1997. An epidemiological model
685 incorporating vector population dynamics applied to African cassava mosaic virus
686 disease. *Journal of Applied Ecology* 34:793-806.
- 687 Hughes, G., and Madden, L. V. 2002. Some methods for eliciting expert knowledge of plant
688 disease epidemics and their application in cluster sampling for disease incidence. *Crop*
689 *Protection* 21:203-215.
- 690 Jones, R. A. C. 2006. Control of plant virus diseases. *Advances in Virus Research* 67:205-244.
- 691 Jones, R. A. C. 2016. Future scenarios for plant virus pathogens as climate change progresses.
692 Pages 87-147 in: *Advances in Virus Research, Vol 95*, M. Kielian, K. Maramorosch and
693 T. C. Mettenleiter, eds.
- 694 Knol, A. B., Slottje, P., van der Sluijs, J. P., and Lebret, E. 2010. The use of expert elicitation in
695 environmental health impact assessment: a seven step procedure. *Environmental Health*
696 9:19.
- 697 Kreuze, J. F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., and Simon, R. 2009.
698 Complete viral genome sequence and discovery of novel viruses by deep sequencing of
699 small RNAs: A generic method for diagnosis, discovery and sequencing of viruses.
700 *Virology* 388:1-7.
- 701 Lamichhane, J. R., Barzman, M., Booi, K., Boonekamp, P., Desneux, N., Huber, L., Kudsk, P.,
702 Langrell, S. R. H., Ratnadass, A., Ricci, P., Sarah, J.-L., and Messean, A. 2015. Robust
703 cropping systems to tackle pests under climate change. A review. *Agronomy for*
704 *Sustainable Development* 35:443-459.
- 705 Legg, J. 1999. Emergence, spread and strategies for controlling the pandemic of cassava mosaic
706 virus disease in east and central Africa. *Crop Protection* 18:627-637.
- 707 Legg, J. P., Kumar, P. L., Makesh Kumar, T., Tripathi, L., Ferguson, M., Kanju, E.,
708 Ntawuruhunga, P., and Cuellar, W. 2015. Cassava virus diseases: Biology, epidemiology,
709 and management. *Advances in Virus Research* 91:85-142.
- 710 Legg, J. P., Jeremiah, S. C., Obiero, H. M., Maruthi, M. N., Ndyetabula, I., Okao-Okuja, G.,
711 Bouwmeester, H., Bigirimana, S., Tata-Hangy, W., Gashaka, G., Mkamillo, G., Alicai, T.,
712 and Kumar, P. L. 2011. Comparing the regional epidemiology of the cassava mosaic and
713 cassava brown streak virus pandemics in Africa. *Virus Research* 159:161-170.

- 714 Levins, R. 1966. The strategy of model building in population biology. *American Scientist*
715 54:421-431.
- 716 Mac Nally, R. 2007. Consensus weightings of evidence for inferring breeding success in broad-
717 scale bird studies. *Austral Ecology* 32:479-484.
- 718 Madden, L. V., Hughes, G., and Van Den Bosch, F. 2007. The study of plant disease epidemics.
719 *American Phytopathological Society St. Paul, MN.*
- 720 Mallowa, S. O., Isutsa, D. K., Kamau, A. W., and Legg, J. P. 2011. Effectiveness of
721 phytosanitation in cassava mosaic disease management in a post-epidemic Area of
722 Western Kenya. *ARPN Journal of Agricultural and Biological Science* 6:8-15.
- 723 Mallowa, S. O., Isutsa, D. K., Kamau, A. W., Obonyo, R., and Legg, J. P. 2006. Current
724 characteristics of cassava mosaic disease in postepidemic areas increase the range of
725 possible management options. *Annals of Applied Biology* 149:137-144.
- 726 Martin, T. G., Kuhnert, P. M., Mengersen, K., and Possingham, H. P. 2005. The power of expert
727 opinion in ecological models using Bayesian methods: Impact of grazing on birds.
728 *Ecological Applications* 15:266-280.
- 729 McGuire, S., and Sperling, L. 2016. Seed systems smallholder farmers use. *Food Security* 8:179-
730 195.
- 731 McQuaid, C. F., Sseruwagi, P., Pariyo, A., and van den Bosch, F. 2016. Cassava brown streak
732 disease and the sustainability of a clean seed system. *Plant Pathology* 65:299-309.
- 733 Moslonka-Lefebvre, M., Finley, A., Dorigatti, I., Dehnen-Schmutz, K., Harwood, T., Jeger, M.
734 J., Xu, X., Holdenrieder, O., and Pautasso, M. 2011. Networks in plant epidemiology:
735 From genes to landscapes, countries, and continents. *Phytopathology* 101:392-403.
- 736 Mwangi, J. K., Nyende, A. B., Demo, P., and Matiru, V. N. 2008. Detection of latent infection
737 by *Ralstonia solanacearum* in potato (*Solanum tuberosum*) using stems instead of tubers.
738 *African Journal of Biotechnology* 7:1644-1649.
- 739 Oberkampf, W. L., Helton, J. C., Joslyn, C. A., Wojtkiewicz, S. F., and Ferson, S. 2004.
740 Challenge problems: uncertainty in system response given uncertain parameters.
741 *Reliability Engineering & System Safety* 85:11-19.
- 742 Parsa, S., Morse, S., Bonifacio, A., Chancellor, T. C. B., Condori, B., Crespo-Perez, V., Hobbs,
743 S. L. A., Kroschel, J., Ba, M. N., Rebaudo, F., Sherwood, S. G., Vanek, S. J., Faye, E.,
744 Herrera, M. A., and Dangles, O. 2014. Obstacles to integrated pest management adoption
745 in developing countries. *Proceedings of the National Academy of Sciences, USA*
746 111:3889-3894.
- 747 R Core Team. 2015. R: A language and environment for statistical computing. R foundation for
748 statistical computing.
- 749 Rahman, M. S., and Akanda, A. M. 2008. Impact of PVY and PLRV on growth and yield of
750 second generation seed potato. *Bangladesh Journal of Plant Pathology* 24:19-24.
- 751 Restif, O., Hayman, D. T. S., Pulliam, J. R. C., Plowright, R. K., George, D. B., Luis, A. D.,
752 Cunningham, A. A., Bowen, R. A., Fooks, A. R., O'Shea, T. J., Wood, J. L. N., and
753 Webb, C. T. 2012. Model-guided fieldwork: practical guidelines for multidisciplinary
754 research on wildlife ecological and epidemiological dynamics. *Ecology Letters* 15:1083-
755 1094.
- 756 Salazar, L. F. 1996. Potato viruses and their control. International Potato Center, Lima, Peru.
- 757 Schulte-Geldermann, E., Gildemacher, P. R., and Struik, P. C. 2012. Improving seed health and
758 seed performance by positive selection in three Kenyan potato varieties. *American*
759 *Journal of Potato Research* 89:429-437.

- 760 Shaw, M. W., and Pautasso, M. 2014. Networks and plant disease management: Concepts and
761 applications. *Annual Review of Phytopathology* 52:477-493.
- 762 Sisterson, M. S., and Stenger, D. C. 2013. Roguing with replacement in perennial crops:
763 Conditions for successful disease management. *Phytopathology* 103:117-128.
- 764 Thiele, G. 1999. Informal potato seed systems in the Andes: Why are they important and what
765 should we do with them? *World Development* 27:83-99.
- 766 Thomas-Sharma, S., Abdurahman, A., Ali, S., Andrade-Piedra, J., Bao, S., Charkowski, A.,
767 Crook, D., Kadian, M., Kromann, P., and Struik, P. 2016. Seed degeneration in potato:
768 the need for an integrated seed health strategy to mitigate the problem in developing
769 countries. *Plant Pathology* 65:3-16.
- 770 van den Bosch, F., Jeger, M. J., and Gilligan, C. A. 2007. Disease control and its selection for
771 damaging plant virus strains in vegetatively propagated staple food crops; a theoretical
772 assessment. *Proceedings of the Royal Society B-Biological Sciences* 274:11-18.
- 773 van der Waals, J. E., Kruger, K., Franke, A. C., Haverkort, A. J., and Steyn, J. M. 2013. Climate
774 change and potato production in contrasting South African agro-ecosystems 3. Effects on
775 relative development rates of selected pathogens and pests. *Potato Research* 56:67-84.
- 776 Winkler, R. L. 1996. Uncertainty in probabilistic risk assessment. *Reliability Engineering &*
777 *System Safety* 54:127-132.
- 778 Yamada, K., Elith, J., McCarthy, M., and Zenger, A. 2003. Eliciting and integrating expert
779 knowledge for wildlife habitat modelling. *Ecological Modelling* 165:251-264.
- 780

781 **Tables & Figures**

782

783 Table 1. State variables monitored in seed degeneration risk assessment framework.

State Variable	Description
HP_s	Number of healthy plants in field, end of season
DP_s	Number of diseased plants in field, end of season
HS_s	Number of healthy seed, produced end of season
DS_s	Number of diseased seed, produced end of season
pHS_s	Proportion of healthy seed, used in following season
pDS_s	Proportion of diseased seed, used in following season
Y_s	Absolute units of yield, end of season
YL_s	Percentage yield loss, end of season

784

785 Table 2. Parameters used in seed degeneration risk assessment framework.

Parameter	Description	Biological meaning of values	Default values used
pHS_0	Initial proportion of healthy seed	1=no seed infected 0=all seed infected	0.8 (low starting infection scenarios) 0.2 (high starting infection scenarios)
K	Initial plant population (number)	Population at beginning of season based on planting rate in a small field	100
E	External inoculum	Amount of host/non-host inoculum surrounding a field	0 (absence of external inoculum) 30 (presence of external inoculum)
β	Maximum transmission rate per season	Maximum rate of disease transmission during the season when there are no limiting factors for disease spread	0.02
W	Proportional change in infection due to environment	W=1, maximally conducive environmental conditions W=0, environmental conditions that do not support transmission	0.8 (highly disease-conductive weather) 0.2 (marginally disease-conductive weather)
H^1	Proportional change in infection due to host genetic resistance	H=1, highly susceptible H=0, immune	0
M^1	Proportional change in infection rate due to vector management	M=1, indicates no management M=0, indicates vector or pathogen eradication	0
A^1	Proportion diseased plants remaining after roguing	A=1, indicates no roguing A=0, indicates all diseased plants removed	0
G	Seed production rate in healthy plants	Number of seed produced per healthy plant	4

Z^1	Proportional selection against diseased plants (through positive or negative selection)	$Z=1$, indicates no seed selection $Z<1$, indicates proportional selection against diseased plants $Z=0$, indicates complete selection against diseased plants	0
C	Indicates differential seed production in the diseased plants as a proportion of seed production in healthy plants	$C=0$, indicates no seed production in diseased plants $C=1$, indicates no difference in seed production between healthy and diseased plants $C<1$, indicates reduced seed production in diseased plants $C>1$, indicates increased seed production in diseased plants	0.9
R	Reversion rate	Proportion of diseased plants that produce disease-free seed	0.1
Φ	Proportion certified (or otherwise completely disease-free) seed purchased	$\phi=1$, all certified seed $\phi=0$, no certified seed	0
Θ	Rate of decline of end of season yield with increasing disease incidence	$0<\theta\leq 0.5$, indicates yield decline slow initially, then increases $\theta=\text{negative}$, indicates yield decline is rapid initially, then slows $\theta=0$, indicates constant rate of decline	0.2
γ	Proportional change in effect of disease incidence on yield loss for late season versus early season	$\gamma=0$, indicates no yield loss due to late season disease incidence $\gamma=1$, indicates no difference between early and late season effects of disease incidence on yield loss	Not used in general models
minY	Minimum yield	Units of yield produced by a severely infected plant	0
maxY	Maximum yield	Units of yield produced by a healthy plant	100

786 ¹ When addressing results, we describe and discuss these management effects in terms of the

787 effectiveness of management implementation, so that all types of management can be considered

788 with 1 indicating complete effectiveness of implementation and 0 indicating complete
789 ineffectiveness. In contrast, for H, M, A, and Z, the model and code are constructed such that 1
790 indicates no limiting factor for infection processes.

791 Table 3. Minimum effectiveness of management practices required to keep average yield loss
 792 below 10% after 10 seasons, under different combinations of weather conduciveness, host
 793 resistance, and low starting levels of infection, in the absence of external inoculum (based on
 794 2000 simulations). Host resistance is expressed as 0 for the highly susceptible host and 0.6 for a
 795 moderately resistant host. For each of the other management components tested, a range of
 796 values from 0 to 1 at increments of 0.1 was evaluated. Other parameters had default values from
 797 Table 2.

Management practice	Highly disease-conducive weather		Marginally disease-conducive weather		Interpretation
	Susceptible host	Resistant host	Susceptible host	Resistant host	
Vector or pathogen management	0.9	0.7	0.7	0.0	Proportional effectiveness of vector or pathogen management
Roguing	0.7	0.4	0.3	0.0	Proportion diseased plants removed
Seed selection	0.7	0.4	0.3	0.0	Proportion healthy seed selected
Use of certified seed	0.6	0.3	0.3	0.1	Proportion certified seed used for planting

798

799 Table 4. Mean number of seasons until proportion of healthy seed falls below 0.7, in the absence
 800 of external inoculum, when a maximum of 15 seasons are considered (based on 2000
 801 simulations). Host resistance and seed selection were evaluated at 0 or 0.6 proportional
 802 effectiveness of management. Other parameters had the default values from Table 2.

Proportional effectiveness of management		Highly disease-conducive weather		Marginally disease-conducive weather	
Seed selection	Host resistance	Absence of external inoculum	Presence of external inoculum	Absence of external inoculum	Presence of external inoculum
0	0	1.2	1.0	6.1	2.1
0.6	0	11.6	1.2	>15.0	8.6
0	0.6	3.1	1.2	14.9	5.6
0.6	0.6	>15.0	11.6	>15.0	>15.0

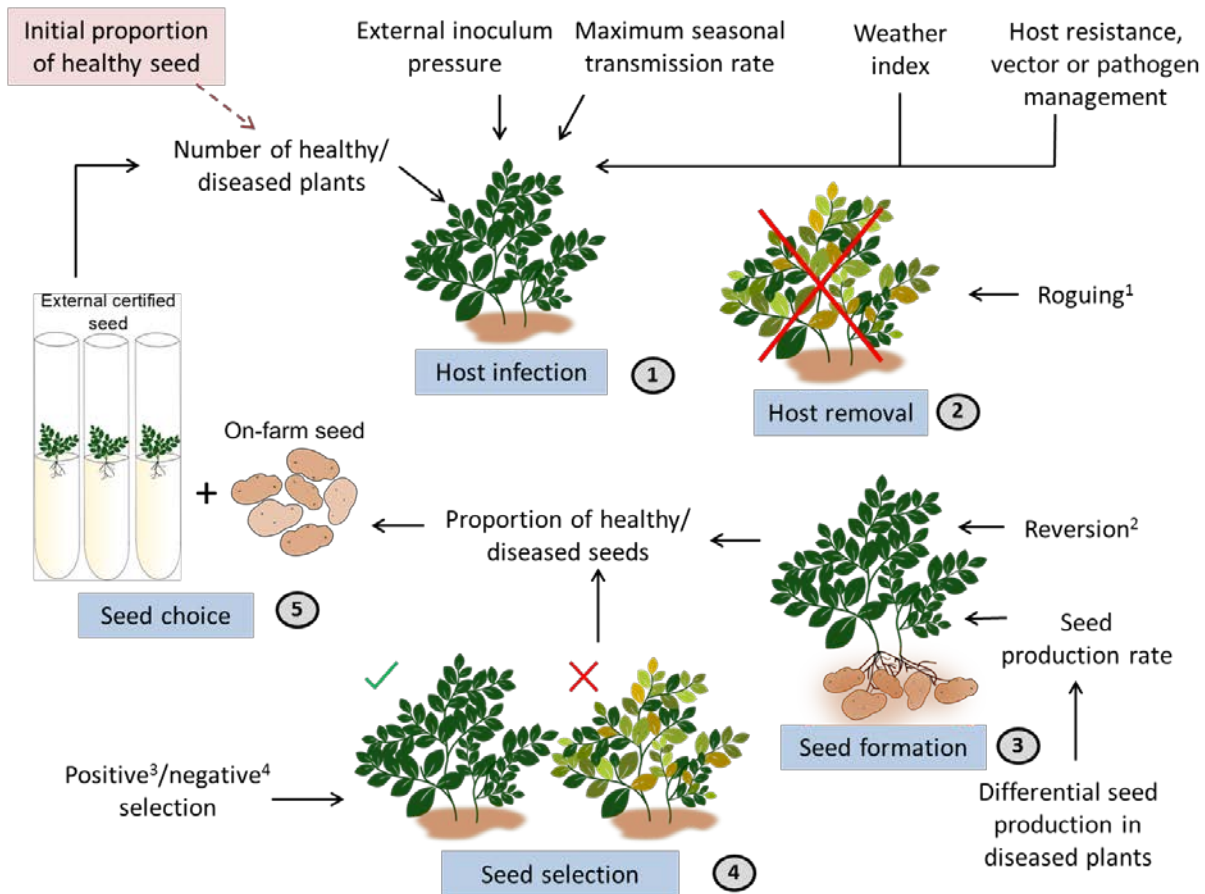
803

804 Table 5. Yield loss incurred for average, near worst-case, and near best-case outcomes when
 805 seed selection and host resistance are used at 0.6 effectiveness of implementation, under highly
 806 disease-conducive weather conditions (based on 2000 simulations). Host resistance and seed
 807 selection are expressed as the proportion effectiveness of management implementation (1
 808 indicating complete effectiveness, and 0 indicating no management). Other parameters had the
 809 default values in Table 2.

Proportional effectiveness of management		Absence of external inoculum			Presence of external inoculum		
Seed selection	Host resistance	5 th percentile (Near best-case outcome)	Mean outcome	95 th percentile (Near worst-case outcome)	5 th percentile (Near best-case outcome)	Mean outcome	95 th percentile (Near worst-case outcome)
0	0	89	98	100	95	99	100
0.6	0	5	31	86	84	98	100
0	0.6	25	42	56	65	85	97
0.6	0.6	1	2	4	16	31	50

810

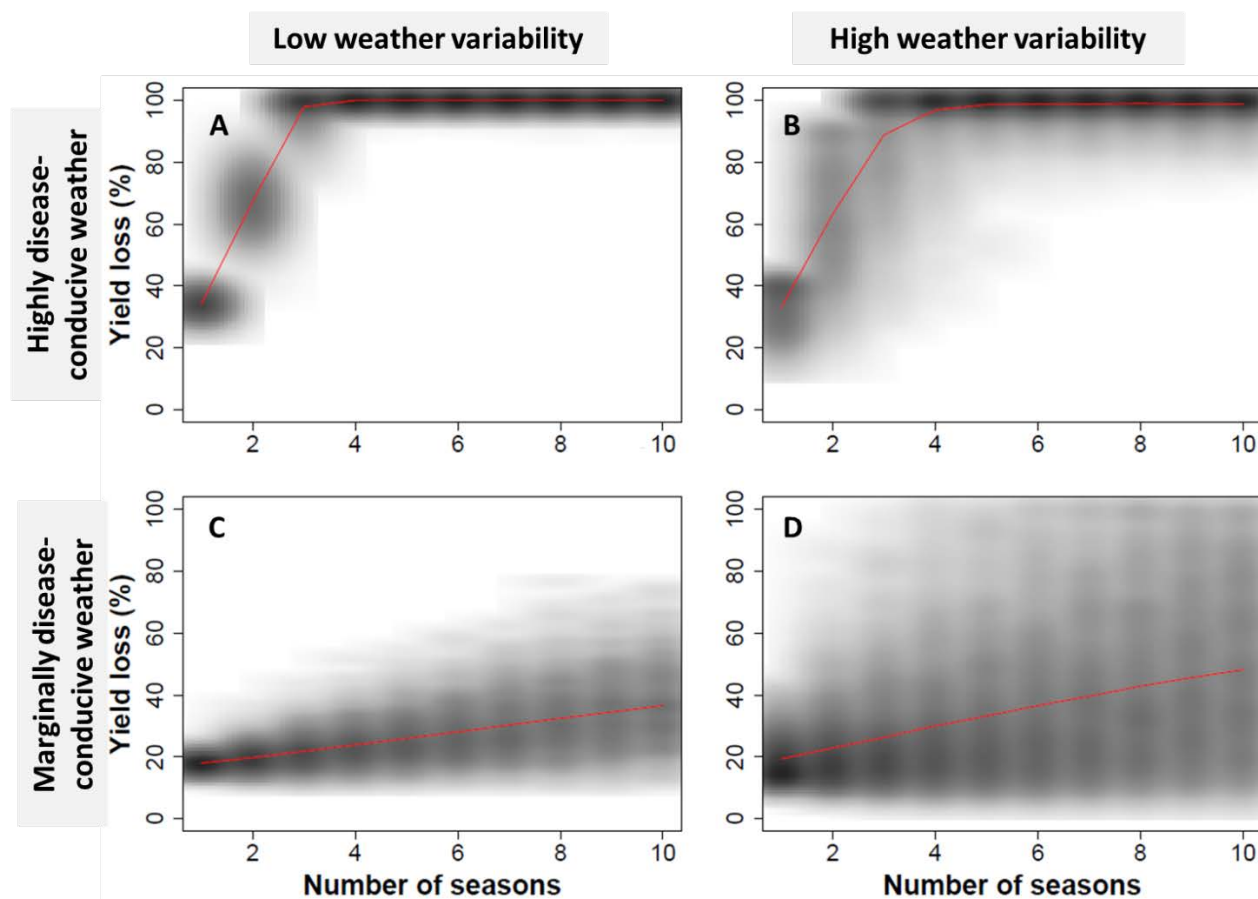
811



812

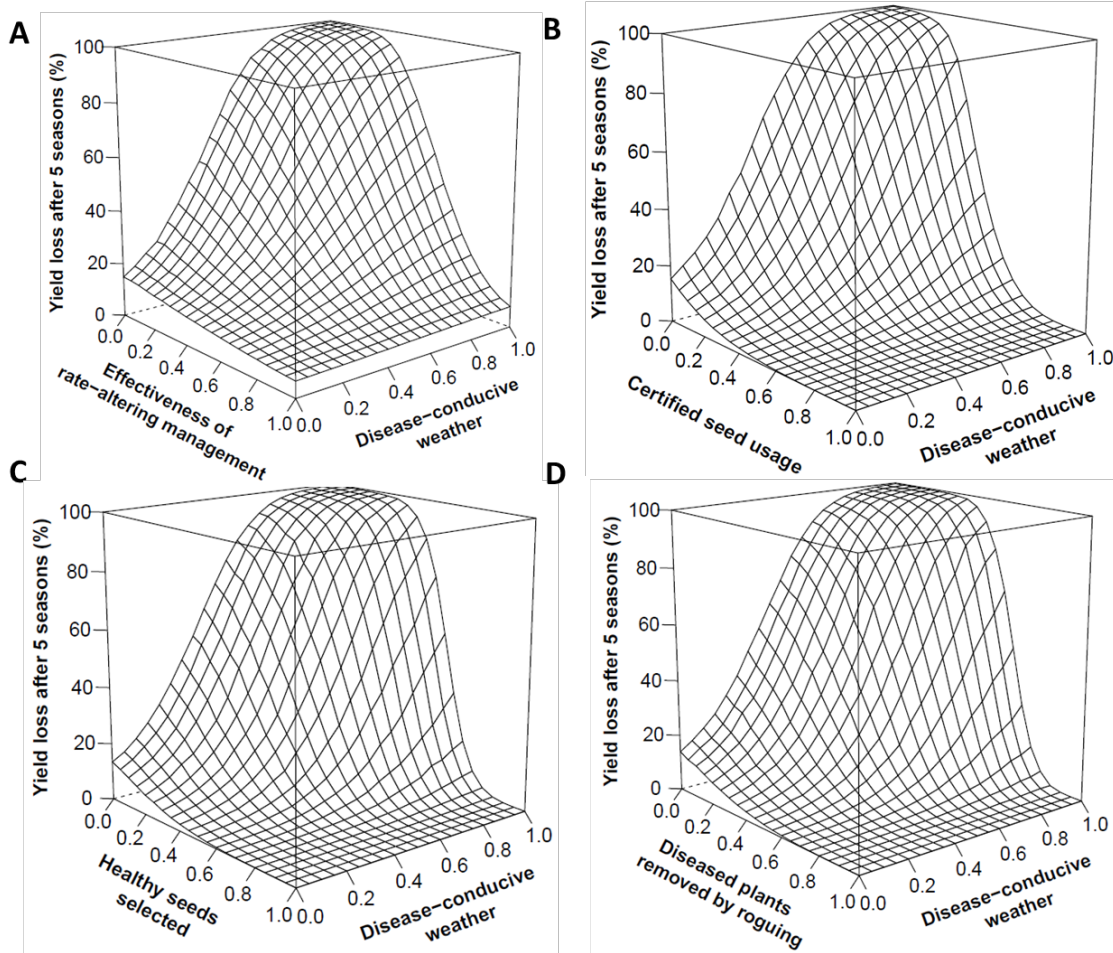
813 Fig.1. Processes modeled in the seed degeneration risk assessment framework (with potato as an
 814 example) are host infection (disease transmission), host removal, seed formation, seed selection,
 815 and seed choice. Rate-altering management components (host resistance, vector or pathogen
 816 management), incidence-altering management components (seed selection, certified seed usage,
 817 and roguing), and phenomena such as reversion and differential seed production in diseased
 818 plants modify these processes. The rate of disease transmission is determined by disease-
 819 conducive weather conditions (included as a ‘weather index’), external inoculum present, the
 820 maximum seasonal transmission rate and any rate-altering management components. In the first
 821 season (time-step), the initial proportion of healthy seed used is provided, after which other
 822 processes and phenomena are introduced in the order depicted by the circled numbers.

- 823 ¹Removal of diseased plants from the field
- 824 ²Production of disease-free seed by infected mother plant
- 825 ³Selection of asymptomatic plants for seed under high disease intensity
- 826 ⁴Rejection of symptomatic plants for seed under low disease intensity



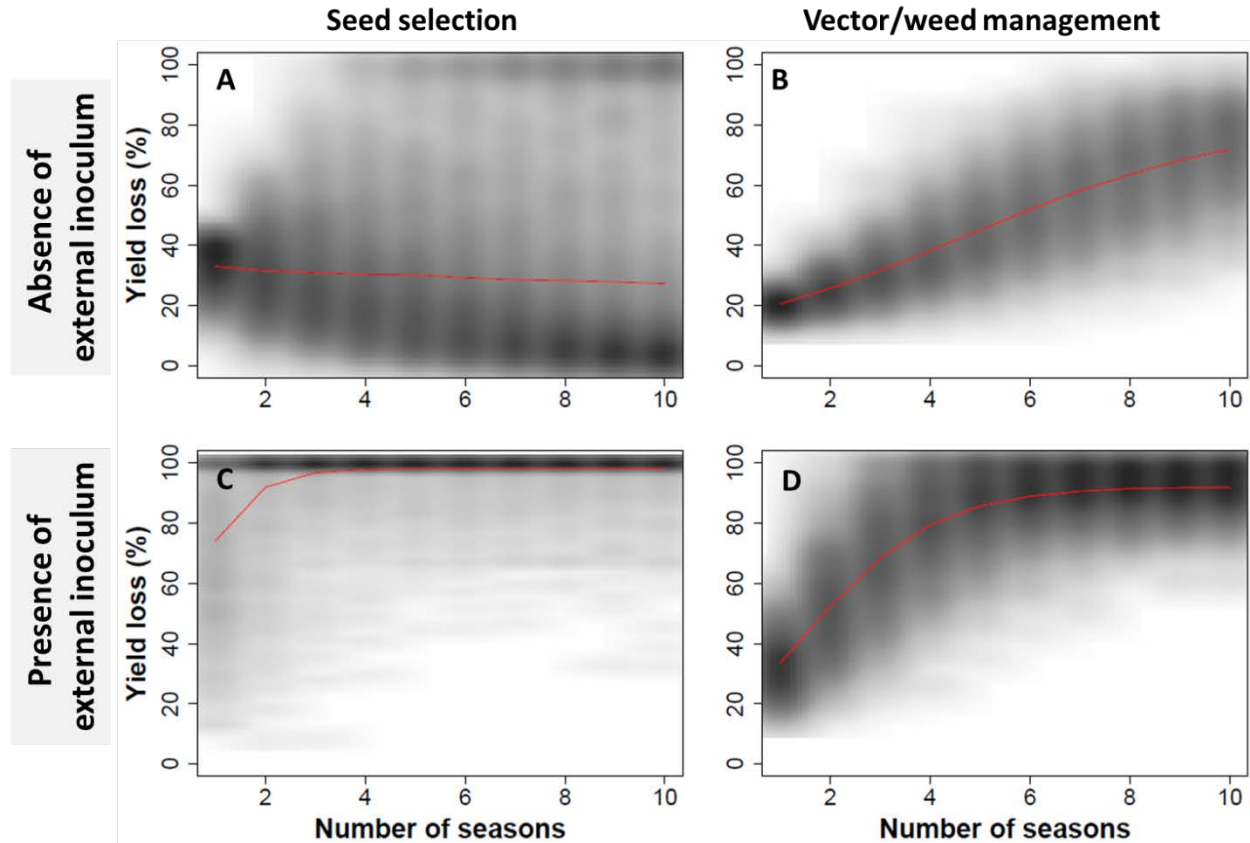
827

828 Fig. 2. Long-term (10 season) yield loss under no-management scenarios, with low starting
829 levels of infection, under highly and marginally disease-conductive weather scenarios with high
830 (0.3) and low (0.1) season-to-season variability in weather, in the absence of external inoculum
831 (based on 2000 simulations). Other parameters are set to default values from Table 2. Red lines
832 indicate the mean value.



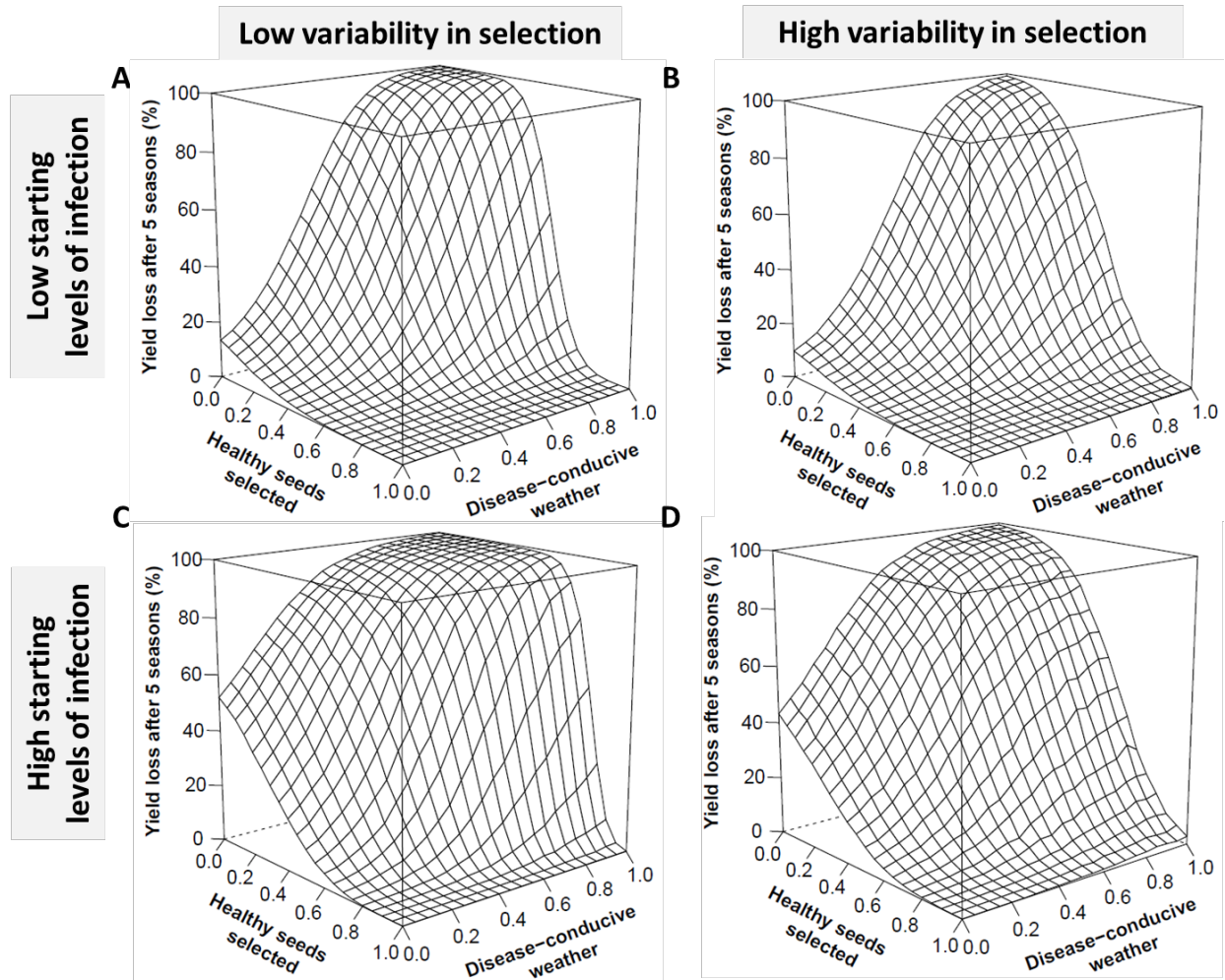
833

834 Fig. 3. Effect of rate-altering management components (such as vector or pathogen management
835 and host resistance) (A) and incidence-altering management components—certified seed usage
836 (B), seed selection (C), and roguing (D) on percent yield loss after 5 seasons under varying
837 disease-conductive weather conditions, high variability in weather (0.3) and low starting levels of
838 infection, in the absence of external inoculum (based on 2000 simulations). Vector management,
839 seed selection, and roguing assume low variability in effectiveness and are expressed as the
840 proportion effectiveness of management implementation (1 indicating complete effectiveness,
841 and 0 indicating no management). Other parameters are default values from Table 2.



842

843 Fig. 4. Long-term (10 season) yield loss under seed selection and vector or pathogen
844 management, in the presence and absence of external inoculum (based on 2000 simulations).
845 Vector management and seed selection, expressed in terms of the proportion effectiveness of
846 implementation, had low variability (0.1) and were each set to 0.6 effectiveness of
847 implementation. Other parameters are default values from Table 2. Red line indicates the mean
848 value.



849

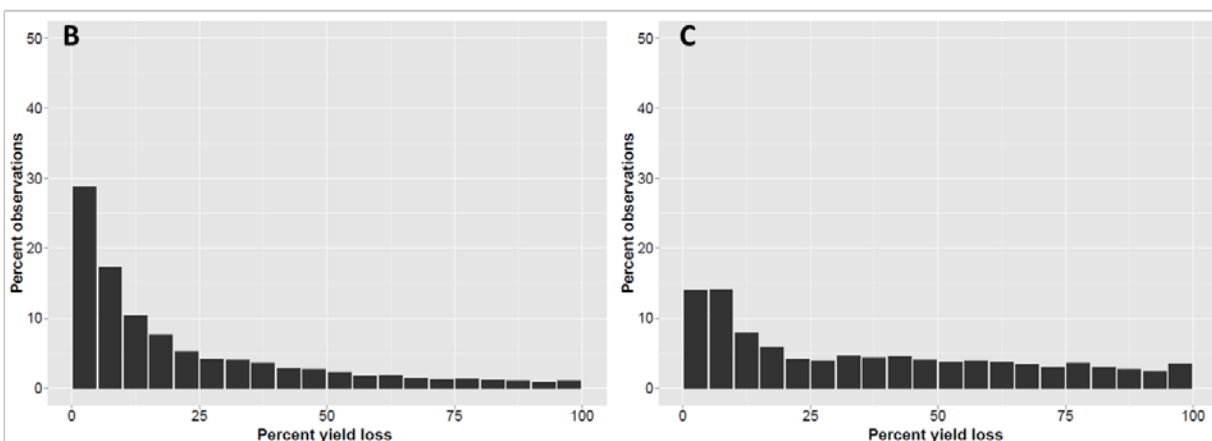
850 Fig. 5. Percent yield loss after 5 seasons as a function of the mean effectiveness of seed selection
851 (proportion healthy seeds selected) for the range of potential levels of disease-conduciveness of
852 weather, at low (A, B) and high (C, D) starting levels of infection, and low (0.1; A, C) and high
853 (0.3; B, D) variability in selection, in the absence of external inoculum (based on 2000
854 simulations). Seed selection is expressed in terms of the effectiveness of implementation. Other
855 parameters are default values from Table 2.

A Mean of experts' estimate of cassava field acreage under different resistance levels

	0 = varieties grown are susceptible					100 = varieties grown are immune				
Resistance levels	0 – 10	11 – 20	21 – 30	31 – 40	41 – 50	51 – 60	61 – 70	71 – 80	81 – 90	91 – 100
Field acreage	38.75	22.5	0	5	3.75	0	21.25	0	7.5	1.25

No information on resistance deployment

With expert elicited information



856

857 Fig. 6. An example of the impact of including expert opinion about the frequency distribution of

858 resistance deployment, compared to an analysis with no information about resistance

859 deployment. A. Mean of experts' estimate of cassava field acreage under different resistance

860 levels. B. Distribution of yield loss in the absence of information on resistance deployment (i.e.,

861 all resistance levels are considered equally likely). C. Distribution of yield loss when experts'

862 estimate is used to weight resistance levels. The other parameters used in the illustration are

863 $pHS_0=0.8$, $K=100$, $E=0$, $\beta=0.02$, $W=0.3$ (high variability [0.3]), $M=1$ (low variability [0.1]), $A=1$

864 (low variability [0.1]), $G=4$, $Z=1$ (low variability [0.1]), $Y=0.9$, $R=0.1$, $\phi=0$, $\theta=0.2$).