Preprint 11-June-2003: Encyclopedia of Life Support Systems (EOLSS)[1]
Section 1.5 "Land Use and Land Cover"
Topic 1.5.27 "Biophysical Models in Land Evaluation"

BIOPHYSICAL MODELS IN LAND EVALUATION

D. G. Rossiter, International Institute for Geo-Information Science and Earth Observation (ITC), Enschede, the Netherlands[2]

Key-words

Land Evaluation, Modelling, Biophysical Land Suitability, Dynamic Simulation, Empirical, Mechanistic, crop yield, Land Qualities, Calibration, Validation

Contents

---

[1] **Complete EOLSS available on CD-ROM and WWW from http://www.eolss.com**

[2] e-mail: rossiter@itc.nl; web page: http://www.itc.nl/personal/rossiter

Glossary

Dynamic simulation model: A model where *time* is explicitly included in the model formulation. The model predicts system state over time, and is driven by a time series of input data, usually weather. Compare *static* model.
Empirical model: A model that relates predictions to data based on previous experience, with no attempt to model the physical causes. Compare *mechanistic* model.
Expert model: A model that attempts to formalise the knowledge of a domain expert, either a land user or, more typically, a specialist in a particular land use.
Mechanistic model: A model that attempts to represent the physical causes of responses to conditions. Compare *empirical* model.
Model: A simplified representation of a *system*, usually in mathematical or computable form.
Model calibration: Adjustment of model *parameters* so that model predictions are close to observed values. Requires a *calibration data set*. Improves the *postdicitive* power of the model.
Model data: Input *variables* to the model that cause *state changes* in the model. They *drive* the behaviour of the model in a particular execution. Examples are daily precipitation (a *time series*) and soil moisture retention by layer (a *static* datum).
Model parameters: Numbers used in the model that are *constants* during a single execution, but which may be variable between executions. They describe the *static* context in which the model is being run. They *parameterize* the equations of the model, i.e., supply specific values that control their numerical behavior. The *form* (equations or algorithm) of the model does not change.
Model validation: Degree of statistical agreement between predictions of the *calibrated* model and observed values. Requires a *validation data set*. Often measured by comparison of observed vs. predicted to a 1:1 line with zero intercept. Quantifies the *predictive* power of the model.
Postdiction: An outcome of modelling a scenario that has already occurred and for which the outcome is known. Compare *prediction*.
Prediction: An outcome of modelling a scenario that has not yet occurred. May be tested in the modelled system by reproducing the input conditions; this is not possible if a specific time series of weather is one of the inputs. Compare *postdiction*.
Simulation: The art of building mathematical or computer models of a system and using these models to study the properties of the system in response to different scenarios.
State variables: Variables whose values represent the state of a system at any time during the simulation. Examples are soil water content by layer and plant biomass.
Static model: A model that does not depend on a time series of input data. Compare *dynamic* model.
System: A limited part of reality, with the connections between its elements and its limits with the outside world (non-system) well-specified.

Summary

Biophysical models are simplified representation of land use systems that allow prediction of the success of such systems prior to their actual implementation. They are classified according to their degree of computation (qualitative to quantitative), descriptive complexity, (empirical to mechanistic) and level in the organizational hierarchy (scale). The simplest models are holistic local knowledge, which is difficult to formalise and can not be extrapolated. Expert models are formalisations of expert judgement about individual Land Qualities, following the FAO Framework for Land Evaluation. Empirical-statistical models are quantitative predictions of crop yield from a set of static Land Characteristics. Dynamic simulation models attempt to model biophysical mechanisms, based on the laws of nature, to follow a system over time based on a time series of input data. Widely applied models in these various categories are discussed here, including ALES, MicroLEIS, WOFOST, PS123, DSSAT, APSIM, EPIC, GAPS, and LEACHM. A stepwise approach is recommended, with simpler models being applied to limit the areas in which the more complicated models must be calibrated.

1 Introduction

A *model* is a simplified representation of reality with which we can compute outcomes without having to perform actual experiments. In land evaluation, models are computer programs that predict the performance of a land use on a land area, when given information on that area's land characteristics. *Biophysical* models predict the behaviour of the land use system in physical terms such as crop yields, environmental effects, and effect on management. They thus provide a quantified procedure to match land with various actual and proposed land uses, as proposed by the FAO Framework for Land Evaluation. Models can be used to predict *crop yields* under different management strategies, as well as individual *land qualities* that are important components of yield, such as moisture supply, nutrient supply, and radiation balance. They can also be used to evaluate individual land qualities important for the land use but not directly affecting yield, such as erosion hazard, trafficability, and workability.

2 Classification of biophysical models

In 1992, Hoosbeek and Bryant proposed a classification of models of pedogenesis (soil formation), which was adapted by Bouma for land evaluation models (Figure 1). In this scheme, models are classified in three dimensions.

The first two dimensions are shown in Figure 1 on a horizontal plane: (1) the *degree of computation*, ranging from qualitative to quantitative; and (2) the *descriptive complexity*, ranging from empirical to mechanistic. The *degree of computation* refers to the precision of the model's prediction. For example, the simplest qualitative model (at the left of the plane) could predict land suitability as "suitable" or "not suitable", in other words, the use will succeed (to some degree) or fail; this could be adequate for some decisions. The most quantitative model (at the right of the plane) would give precise numerical predictions of crop yields and environmental effects. The *descriptive complexity* refers to the detail with which processes are made explicit in the model. An *empirical* model (at the back edge of the plane) is a model where processes are not known, but where relations are established based on experience. By contrast, a *mechanistic* model (at the front edge of the plane) is a model where processes, not just relations, are modelled.

The third dimension is shown on Figure 1 as the vertical axis passing through the plane formed by the first two dimensions: (3) the *level in the organizational hierarchy* (scale of processes being modelled), which for land evaluation range from region through field and "point" to soil horizons and finally molecular interactions. At any scale level, the first two dimensions are possible; in practice the more quantitative and functional models are generally found at smaller scales.

Along the plane formed by dimensions (1) and (2), Hoosbeek and Bryant distinguished several levels of knowledge, which they termed K1 (user expertise), K2 (expert knowledge), K3 (generalized holistic models), K4 (complex holistic models), and K5 (complex models of system components), which of course grade into each other in any actual model.

K1 models are empirical, qualitative expressions of the land user's experience. These have low descriptive complexity and require no computation. They are applied intuitively within the geographical and phenomenological area of the user's experience. K1 models are difficult to formalise, since they draw on the user's holistic experience, rather than a reductionist problem analysis.

K2 models are also qualitative, but consider mechanisms. In particular, the FAO approach with its analysis of land suitability as a set of Land Qualities has the reductionist structure required for these models, which are built by specialists who are trained to search for causes.

K3 models are empirical but quantitative. These are statistical relations between output (e.g. yield) and input (e.g. precipitation, heat units, soil fertility), usually established by regression analysis on large datasets. Predictive variables are selected based on a reductionist concept of causative factors. They can not be applied outside their area of calibration. All variables are static, and there is no attempt to simulate system behaviour over time. They can only be applied to LUTs that are widely practised, so are not useful for new crops, new technologies, or new management strategies.

K4, and K5 models attempt to be mechanistic rather than empirical. This means that they are based more on scientific principles (laws such as conservation of mass and energy, diffusion, convection and dispersion, chemical kinetics and equilibrium) and less on site-specific empirical relations. It is thus expected that they will be 'universally' applicable. However, the line between empirical and mechanistic models is not clear, since all 'mechanistic' models have empirical components. These models, when applied to land evaluation, are usually driven by *daily weather data*. This allows the analysis of *dynamic* and *transient* phenomena that may affect land performance, so that these are commonly referred to as *dynamic simulation models*. Such models can be used to model individual *land qualities* such as moisture sufficiency (K5) as well as crop yield (K3). This is appropriate if the *timing* of the quality is important. Water stress is a good example: the yearly moisture deficit often isn't as important as the deficit in specific parts of the crop growth cycle.

We now consider these modelling approaches, from least to most sophisticated.

3 Models of expert knowledge

At the time that the FAO Framework was developed (mid 1970's), K2 expert knowledge was captured as a set of *matching tables*, one for each Land Quality, using the *maximum limitation method*, requiring a set of *diagnostic Land Characteristics* as input for each table. This was put in computable form and at the same time made more flexible by the ALES ('Automated Land Evaluation System') computer program, which was released in 1986 and improved until 1997. It is freely distributed by Cornell University, but requires a code from a commercial database vendor for legal operation.

ALES provides a *framework* with which land evaluators can build their own *expert systems* to evaluate land according to FAO Framework. Models are built to satisfy local needs, so that ALES does not provide a fixed list of LUT, LUR, or LC. Rather, these lists are defined by the expert to suit local conditions and objectives. ALES does not include any knowledge about land and land use; these come from the expert. A good example of an ALES model is the LEV-CET model for central Ethiopia developed by Yizengaw and Verheye.

A key innovation in ALES is the use of *decision trees* instead of maximum limitation tables to infer Land Qualities from a set of diagnostic Land Characteristics. These are hierarchical multi-way keys in which the *leaves* are results (severity levels of the LQ), and the interior nodes (*branch points*) of the tree are decision criteria (LC values). They are constructed by the model builder, and traversed during the computation of an evaluation result, using actual land data for each land evaluation unit.

Figure 2 shows a simple decision tree adapted from the Fertility Capability Classification of Sánchez and Buol (and now incorporated into the FAO's Topsoil Classification). The objective in this case is to predict the soil-related LQ "*risk of P fixation by iron*"; this LQ limits agricultural systems on some highly-weathered soils where the agronomist attempts to compensate for low soil P by moderate fertilization. Some soils fix ('eat') added P in an unavailable form, so that moderate doses are effectively wasted inputs. In the displayed tree, the diagnostic LC at the highest level is "*ratio of free Fe_2O_3 to clay in the topsoil*"; if this is below an expert-defined threshold (< 0.15), there is no risk and the decision is taken. If the ratio is higher, the second-level diagnostic LC "*percentage of clay in the topsoil*" must be considered; if this is below another expert-defined threshold (here, 35%), again there is no risk; above this threshold there is risk and in either case the decision is taken. A third possibility is that one of the two diagnostic tests was not done, perhaps because of expense or unavailability of a laboratory. In both cases, the expert allows the use of alternative LCs: "*hue (basic colour) of the topsoil matrix*", followed in some cases by the "*topsoil structure*", both indicative of the form of iron-dominated clays. Since these two LCs can always be assessed in the field, a decision can always be taken. Note that the choice of LCs, the thresholds, and the decisions, represents *expert judgement*, in this case of an expert on fertilization of these soils.

4 Empirical-statistical models

An attractive option in the case of well-established LUTs is to model their output as a *static* function of a set of Land Characteristics that are expected to influence the output (K3 models). For the output 'crop yield', these characteristics are typically climate (precipitation and temperature, either average, in the growing season, or in specific periods) and soil (nutrient status, reaction, organic matter content, particle-size distribution). The functions are usually developed by multiple regression techniques. Validated relations can be used as an objective basis for land valuation and taxation, since they deal with widely-grown crops. This approach goes back to the 1930's, when soil surveyors were asked to predict crop yields on the basis of soil properties.

A typical example of this approach is the work of Olson and colleagues in New York State, and later in Illinois (USA). He begins with a conceptual model of maize yield as a function of rainfall, temperature, management, site, topography, soil chemical and physical characteristics, mineralogy, and organisms. The management level (in this case, the level of fertilization, liming, pesticide use and tillage) is fixed, while the other factors are quantified by measured values of land characteristics. For example: the 'rainfall' conceptual factor was approximated by the measured variable 'total yearly rainfall', 'temperature' by 'growing degree days, and 'topography' by 'drainage class (depth to redoximorphic mottles)'. Multiple stepwise linear regression was used to develop increasingly-complicated equations, the best of which (including rainfall, soil depth, soil available water capacity, temperature, sum of basic cations, and organic C content) explained two-thirds of the observed variance (calibration $r^2 = 0.66$); this is a good result for such models. Note that this is *not* the predictive success (validation) of this model, which is expected to be lower; see subsection 'calibration vs. validation' for a discussion of this distinction.

This approach has been computerized, for example as the 'Albero' component of the MicroLEIS system, which predicts yields of maize, cotton and wheat from a set of soil characteristics within a fairly homogeneous climate zone (Sevilla province, Spain) using equations developed by multiple regression.

As with any empirical model, these can only be applied in their original zone of calibration; extrapolation to new conditions is not justified. The usual cautions about developing multiple regressions apply: correlated predictors, unknown true functional form, and parsimony.

5 Dynamic simulation models of crop yield

Over the past 25 years, many individual modellers and collaborative groups have attempted to develop models that simulate the growth of crops, along with associated phenomena that influence crop growth such as water and solute movement in soils. In this section we describe WOFOST and its derivative PS123, DSSAT (incorporating CERES and the GRO models), APSIM, EPIC, and GAPS. These are examples of K4 models; however it bears repeating that these supposedly 'mechanistic' models have a large empirical component in their descriptions of sub-systems (lower levels in the scale hierarchy).

5.1 The WOFOST approach

In the early 1980's, the Center for World Food Studies (CABO) in Wageningen (NL) developed a flexible model based on basic plant physiology and soil processes to predict yields under several *production levels*. The original model is known as WOFOST ('WOrld FOod STudies'); it is in version 7.1.2 (April 2003). In the early 1990's, this was adapted for didactic purposes as the PS123 ("Production Situations 1, 2, and 3") model. It was also incorporated into models of farming systems. A major advantage of this approach is that we can begin with simple models of controlled production situations, where we can have high confidence in the model results, and then increase the complexity for less controlled situations, with correspondingly less confidence in the outcome.

WOFOST has been incorporated into a large number of system models as reviewed by van Ittersum *et al*. Since 1990 all the so-called 'Wageningen' models have been programmed in the FORTRAN Simulation Environment (FSE), and so share many components. WOFOST is sold by Alterra at a low price; other models in this family are also sold by Alterra. PS123 is free.

The WOFOST approach, as explained by van Keulen and Wolf, considers three levels of increasingly more realistic limitations, and correspondingly detailed models:

5.1.1 Production level 1: Radiation and temperature limited

Growth occurs in conditions with ample plant nutrients, water, and oxygen (if necessary) all the time. The growth rate of vegetation is determined by weather conditions and the response of the plant to these. This can be approached in practice with very intensively managed irrigated crops. The model is one of photosynthesis, partition of carbohydrate, and physiological growth stages (e.g., flowering, senescence). The only inputs to the model are temperature and radiation (perhaps inferred from cloudiness). There is no need to simulate soil processes at this level; only above-ground physiology is considered.

5.1.2 Production level 2: Water limited

Growth is limited by water shortage at least part of the time, but when sufficient water is available, the growth rate increases up to the maximum rate set by the weather. This can be approached in practice by intensively managed rainfed crops. The model must determine water stress (so, needs to model soil water, the plant root system, and plant transpiration) and its effect on the photosynthetic and growth processes. Another input to the model is precipitation, and the soil profile must be modeled at least for the water balance.

5.1.3 Production level 3: Nitrogen limited

Growth is limited, at least part of the time, by shortage of nitrogen (N) and water or weather at other times. This is usually more limiting than other nutrients because N transformations in the soil are much more rapid than for other nutrients. This is common in rainfed crops even if fertilized according to recommendations and is especially relevant in systems that use animal or green manures. The model must determine soil N dynamics, plant uptake, N use in the plant, and effects of N stress on photosynthesis, partition and growth. Soil temperature can have a large effect on microbial populations, so this must be modeled. Soil organic matter cycling must be modeled.

This approach has been extended to other levels, e.g. nutrient limited (other than N), and pest and disease limited.

5.2 DSSAT

Probably the most widely known and used dynamic simulation models applied to agricultural production are included in what is now termed DSSAT ('Decision Support System for Agro-technology Transfer'), in version 3.5 (stable) and 4 (developmental) as of April 2003. These include the CERES and *GRO models, and indeed DSSAT grew out of previous single-crop models developed in the early 1980's. It has recently been re-structured in version 4 as a set of modules: soil, crop template, weather, and competition. The generic crop model can be parameterized to simulate various crops. The older versions of DSSAT have been used in hundreds of studies of farming systems, and there is widespread experience with its calibration and data requirements.

Since the late 1990's, DSSAT has been distributed by the International Consortium for Agricultural Systems Applications (ICASA) at low cost. Registered users are provided with the program source code on request.

5.3 APSIM

The APSIM ('Agricultural Production Systems Simulator') modelling approach was developed by the Agricultural Production Systems Research Unit (a joint effort between the Queensland State Government and CSIRO) in Australia. It was designed from the start as a set of modules, which may be built by different researchers. It may be applied to both biophysical and economic aspects of the production system. It is especially intended for farm-level decision-making in the face of uncertain weather. APSIM is not sold, rather licensed on a case-by-case basis.

5.4 Others

Several other models can be applied to land evaluation. The EPIC ('Environmental Policy Integrated Climate') model was developed by the USDA to quantify the costs of soil erosion and benefits of soil control in the USA. It includes a simple (and therefore fairly easily parameterized) crop model. Its predictions include grain yield, residues, and soil loss. As a product of the US Government, it is in the public domain, and thus freely available. GAPS ('General-purpose Atmosphere-Plant-Soil Simulator') is a research model developed by Cornell University. It is freely available, with source code and extensive documentation of model procedures. This makes it attractive for learning the principles of dynamic simulation modelling.

6 Dynamic simulation models of individual land qualities

If we only need to model single Land Qualities, specialised models are available that take a more detailed mechanistic approach (K5) than is possible in a holistic model.

A typical application is solute transport in soils, including pesticides and pollutants such as nitrates. The LEACHM model is a good example. It uses basic physio-chemical conceptual models, such as the convection-dispersion equation for chemicals and Richard's equation for soil water redistribution.

Another important Land Quality that is addressed by many models is erosion hazard. Most prominent among these are EUROSEM, RUSLE, EPIC, AGNPS, WEPP (water) and WERS (wind). Some of these, e.g. EPIC, include a simple crop model, mainly to provide estimates of vegetative cover during the growing season and crop residues after.

7 Critical issues in using dynamic simulation models for land evaluation

7.1 Context

Modelling is widely recognized as an art, requiring a mastery of the model structure, of the data requirements, and good judgement to make choices that can greatly affect model results. No model is automatically applicable to all situations, so the analyst must check that the model's context matches that of the target area. This is because all models are simplifications of reality, and simplifications that do not affect results in some geographic or state spaces may be crucial in other spaces. For example, many models incorporating a soil water balance do not account for snowfall or freezing temperatures; these models can not be applied to winter weather in cold climates. As another example, a model incorporating CO_2 exchange through leaves may not take into account the decreasing vapour pressure with increasing altitude, and so can not be applied without modification in highlands. Addiscott expresses this point from the perspective of an experienced modeller: "Because of the element of simplification, no model should be used to make predictions outside the context in which it was developed or beyond the range of the parameter values from which it has been validated."

7.2 Calibration vs. validation

The process of fitting a model, either a regression equation or a dynamic simulation, to observed data is *calibration*, that is, the model parameters are adjusted to best fit the available experiments. In the case of regression, this is part of developing the equation, given a functional form. This yields a goodness-of-fit measure such as r^2, which expresses how well we were able to match the model to the data.

Another name for calibration is *postdiction* (as opposed to *pre*diction, see below), from the Latin 'post' (after) and 'dicere' (to say). This allows the modeller to use the *past* (already observed in experiments or datasets) to make probabilistic statements about the how well the observations are explained by the calibrated model. If the observations were *representative* of the desired sample space, we would expect to obtain the same parameters, within experimental and observational error, in similar repeated studies. However, there is no way to be sure that the calibration sample is indeed representative of the scenario we want to predict.

The correct measure of the predictive success of a model is a *validation* experiment. This uses the *calibrated* model to predict results for a second set of observations, and then compares the observed vs. predicted results. A plot of the validation vs. predictions should lie along a 1:1 line passing through the origin, within error limits; that is, their correlation should be +1 and a fitted regression should have a zero intercept. If these conditions are not met, the original model is not valid. If so, an estimate of the *predictive* power of the model may be obtained from the r^2 of the *validation* regression (not that of the calibration regression).

Another name for validation is *prediction* (as opposed to *post* diction, see above), from the Latin 'prae' (before) and 'dicere' (to say). It allows the modeller to make a statement about how well the model is expected to predict the *future* scenarios.

Clearly, when reporting the success of models, it is validation, not calibration, which is wanted.

7.3 Calibration

Dynamic simulation models have a large number of *parameters* that must be adjusted to the target area, by *calibration*. This poses difficulties because there is no objective way to determine which parameters to adjust, nor even which model output or which points in time to use for calibration. A common approach is to perform a *sensitivity analysis* of model outputs vs. model parameters, and adjust the most sensitive parameters. Since these models have so many parameters, it is possible to fit them quite closely to all but the largest data sets. This is similar to adding higher-order terms to regression models: as explained by Gaugh at a certain point we are fitting noise, not processes. With regression there are statistical methods such as Akaie's information criterion to determine if a model may be over-fitted, but there is no such measure for dynamic models.

7.4 Data

A major impediment to applying dynamic simulation models in routine land evaluation is the requirement for high-quality, high-frequency data. A typical example is the *minimum data set* for the CERES and GRO series of models incorporated into DSSAT (see below). These include:

- Soil properties as a function of depth: horizon thickness, upper and lower limits of volumetric water, volumetric water at saturation, bulk density, pH, organic carbon, total nitrogen;
- Daily weather data: radiation, precipitation, maximum and minimum temperatures;
- Crop parameters: maturity type, photoperiod response, yield components;
- Initial conditions: water content by depth, nitrates and ammonium by depth;
- Management choices: sowing date, plant population, irrigation amounts and dates, fertilizer amounts and dates, residue management, plowing depth.

As can be readily appreciated, these are expensive to obtain, and out of reach, except in research settings, for many land evaluation applications. Especially troublesome are the many parameters that are needed for each crop variety. Daily weather can be approximated from decadal or even monthly data with *weather generator* programs, one of which is included in DSSAT. Missing solar radiation data can be approximated by locally-calibrated transfer functions from cloudiness and latitude. Missing soil properties can be estimated with pedotransfer functions from routine soil survey data. These approximations will typically lead to less successful calibrations.

7.5 Mismatched conceptual levels

Application of models with K5 descriptions of sub-systems often results in a mis-match between land data, provided at the site or point level, and the parameter requirements of the model, provided at the structural aggregate or molecular level. Attempts to bridge this gap usually are by means of empirical statistical models from available land data to model parameters; these have been called ‘pedotransfer functions’. These introduce an additional source of uncertainty into the predictions.

8 Selecting a modelling approach

The biophysical reality that the land evaluator seeks to model is complex, so it is not surprising that modelling is difficult. The FAO Framework provides a clue to selecting appropriate models in a *stepwise* approach. First, areas of non-suitability can be determined by K2 expert-knowledge models, considering important Land Use Requirements, with severity levels of the corresponding Land Qualities being evaluated by decision trees. This is especially relevant for limitations that can not be removed with the chosen technology. For example, citrus can not be grown out of doors where there is hard frost. Second, K3 models can be used for well-established, widely-implemented LUTs where there is sufficient data to develop sound statistical relations with climate and soil characteristics. Third, dynamic simulation models can be used for Land Qualities that depend on dynamics, e.g. moisture sufficiency (K5), and for overall suitability based on crop yield (K4).

Dynamic simulation models have a major advantage compared to static empirical-statistical models if we are trying to understand the processes that contribute to an outcome (e.g. a crop yield), and especially the probable effect of management decisions such as earlier planting of variable levels of fertilisation. However, their complexity, high data requirements, and difficult calibration often make them less reliable in situations where a large number of input and output data are available to statistically fit an empirical model.

No matter which modelling approach is selected, the land evaluator must use judgement and common sense, as well as feedback from clients at every stage in the process, from evaluation objectives, through model selection, to presentation of results.

Bibliography

Addiscott T. M. (1993). Simulation modelling and soil behavior. *Geoderma,* **60**, 15-40. [Excellent overview of the problems in applying models in the context of soil science, including evaluation of soil-related land qualities. Explains problems with non-linearity of models (the average result isn't the result of averages) and discusses calibration vs. validation.]

Bouma J. (1997). The role of quantitative approaches in soil science when interacting with stakeholders (with Discussion). *Geoderma,* **78**, 1-12. [Explains the levels of knowledge (K1 through K5) in land evaluation studies, and illustrates with case studies. Research proceeds by the prototyping of realistic solutions, identification of bottlenecks, and research at the lowest-possible level (closest to the end user of the research). The scientist is seen as a 'knowledge broker']

Bouma J. (1999). Land evaluation for landscape units. In Sumner M. E. (Ed.), *Handbook of soil science* (pp. E393-E412). Boca Raton, FL: CRC Press. [Presents the Hoosbeek & Bryant model classification adapted for land evaluation studies. Also introduces the concept of the "research chain" for demand-driven land evaluation, and defines soil "genoforms" and "phenoforms"]

Driessen P. M. & Konijn N. T. (1992). *Land-use systems analysis*. Wageningen: Wageningen Agricultural University, Department of Soil Science & Geology. [A practical text, using the PS123 adaptation of the WOFOST crop modelling approach.]

Gauch Jr. H. G. (1993). Prediction, parsimony & noise. *American Scientist,* **81**, 468-478. [Discusses postdiction vs. prediction and the benefits of fiitting parsimonious models. Well-illustrated, with clear examples.]

Hutson J. L. & Wagenet R. J. (1991). Simulating nitrogen dynamics in soils using a deterministic model. *Soil Use and Management,* **7**, 74-78. [Explains the concepts behind the LEACHM model of solute movement in soils.]

Ittersum M. K. van, Leffelaar P. A., Keulen H. van, Kropff M. J., Bastiaans L., & Goudriaan J. (2003). On approaches and applications of the Wageningen crop models. *European Journal of Agronomy,* **18**(3-4), 201-234. [An overview of the Wageningen crop and soil modelling approach. Includes detailed descriptions of LINTUL, SUCROS, ORYZA, WOFOST and INTERCOM. Gives examples of applications to plant type design, guiding experimental research, education, yield gap analysis, evaluation of manure policies, crop growth monitoring system and analysis and design of farming and regional land use systems.]

Keating B. A., Carberry P. S., Hammer G. L., Probert M. E., Robertson M. J., Holzworth D., Huth N. I., Hargreaves J. N. G., Meinke H., & Hochman Z. (2003). An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy,* **18**(3-4), 267-288. [Overview of the Agricultural Production Systems Simulator (APSIM) project from the Agricultural Production Systems Research Unit (Queensland State Government and CSIRO) in Australia, A very sophisticated modelling approach, applied to both biophysical and economic aspects of the production system.]

Keulen H. van & Wolf J. (Eds.). (1986). *Modelling of agricultural production: weather, soils and crops*. Wageningen: PUDOC. [A textbook with many worked examples following the WOFOST approach. Includes a chapter by Driessen on soils data for simulation modelling.]

Jones J. W., Hoogenboom G., Porter C. H., Boote K. J., Batchelor W. D., Hunt L. A., Wilkens P. W., Singh U., Gijsman A. J., & Ritchie J. T. (2003). The DSSAT cropping system model. *European Journal of Agronomy,* **18**(3-4), 235-265. [The most recent overview of this influential set of models, which includes CERES and GRO. Includes a long list of DSSAT applications, organised by geographic and application area.]

Olson K. R. & Olson G. W. (1986). Use of multiple regression analysis to estimate average corn yields using selected soils and climatic data. *Agricultural Systems,* **20**, 105-120. [A typical example of multiple regression of crop yields on biophysical land characteristics, in a static K3 model.]

Passioura J. B. (1996). Simulation models: science, snake oil, education or engineering? *Agronomy Journal,* **88**, 690-694. [An experienced modeller discusses the proper use of dynamic simulation models in agronomy and soil science.]

Rossiter D. G. (1996). A theoretical framework for land evaluation (with Discussion). *Geoderma,* **72**, 165-202. [Classifies land evaluation approaches, including simulation modelling. Includes an interesting discussion by experts on trends in land evaluation.]

Rossiter D. G. & Riha S. J. (1999). Modeling plant competition with the GAPS object-oriented dynamic simulation model. *Agronomy Journal,* **91**(5), 773-783. [Explains, and illustrates with pseudo-code, the object-oriented structure of the GAPS dynamic simulation model.]

Wösten J. H. M. (2002). Pedotransfer functions. In Lal R. (Ed.), *The Encyclopedia of soil science* (pp. 967-971). New York: Dekker. [Explains the use of transfer functions from available soil survey data to the parameters required by dynamic simulation models.]

Yizengaw T. & Verheye W. (1995). Application of computer captured knowledge in land evaluation, using ALES in central Ethiopia. *Geoderma,* **66**, 3-4. [This is a typical ALES application, a simple decision procedure, referred to as Land EValuation system for Central EThiopia (LEV-CET), using local expert knowledge to evaluate biophysical suitability for locally-adapted grain crops, considering both climatic and soil factors.]

Websites

Agricultural Production Systems Research Unit. (2003). *APSIM*. Retrieved 31-March, 2003, from http://www.apsru.gov.au/Products/apsim.htm [APSIM home page: system structure, obtaining the model, applications]

Alterra. (2003). *Alterra - Green World Research*. Retrieved 31-March, 2003, from http://www.alterra.dlo.nl/english/ [Source for WOFOST and SWAP models. Click on the 'products' link, and then select the WOFOST model.]

Benz J. (2003, 19-March). *WWW-Server for Ecological Modelling*. Retrieved 27-March-2003, from http://eco.wiz.uni-kassel.de/ecobas.html [A comprehensive list of ecological models, including yield prediction, water relations in soils and catchments, erosion, and pollution. Searchable by model application and structure. Links to each model's home page.]

Hutson J. L. (2002). *Leaching Estimation and Chemistry Model*. Retrieved 02-April, 2003, from http://www.scieng.flinders.edu.au/cpes/people/hutson_j/leachweb.html [Home page for the LEACHM model, used to evaluate land qualities related to soil chemistry and groundwater pollution]

International Consortium for Agricultural Systems Applications. (2003). *International Consortium for Agricultural Systems Applications Home Page*. Retrieved 31-March, 2003, from http://www.icasa.net/ [Source of the DSSAT models. Also promotes decision support systems for agro-technology transfer.]

Meinardus A., Griggs R. H., Verel Benson, & Williams J. (1998, February 18). *Welcome to EPIC*. Retrieved 01-April, 2003, from http://www.brc.tamus.edu/epic/ [Home page of the EPIC (Environmental Policy Integrated Climate) dynamic simulation model.]

Riha S. J. (2003). *Environmental Biophysics and Modeling: Software*. Retrieved 31-March, 2003, from http://environment.eas.cornell.edu/software.htm [Home page of the General-purpose Atmosphere-Plant-Soil Simulator (GAPS) dynamic simulation model.]

de la Rosa D. (2003). *MicroLEIS*. Retrieved 01-April, 2003, from http://www.microleis.com [A set of computer programs for land evaluation by various approaches, especially applicable to the Mediterranean.]

Rossiter D. G. (1995, 10-July-2001). *ALES's Home Page*. Retrieved 02-April, 2003, from http://www.css.cornell.edu/landeval/ales/ales.htm [Documentation and software for the Automated Land Evaluation System (ALES) expert system framework]

Rossiter D. G. (2003, 03-April). *A Compendium of On-Line Soil Survey Information: Models & Applications: Soil Erosion Modelling*. Retrieved 03-April, 2003, from http://www.itc.nl/personal/rossiter/research/rsrch_ss_apps.html#erosion [Links to the home pages of important soil erosion models]

Rossiter D. G. (2003, 03-April). *A Compendium of On-Line Soil Survey Information: Models & Applications: PS123*. Retrieved 03-April, 2003, from http://www.itc.nl/personal/rossiter/research/rsrch_ss_apps.html#ps123 [Download point for the PS123 software and associated documentation]

Figures

Figure 1: Conceptual framework to classify models, after Bouma (1999) Land evaluation for landscape units. In: Sumner M. E. (Ed.), *Handbook of soil science* (pp. E393-E412). Boca Raton, FL: CRC Press.

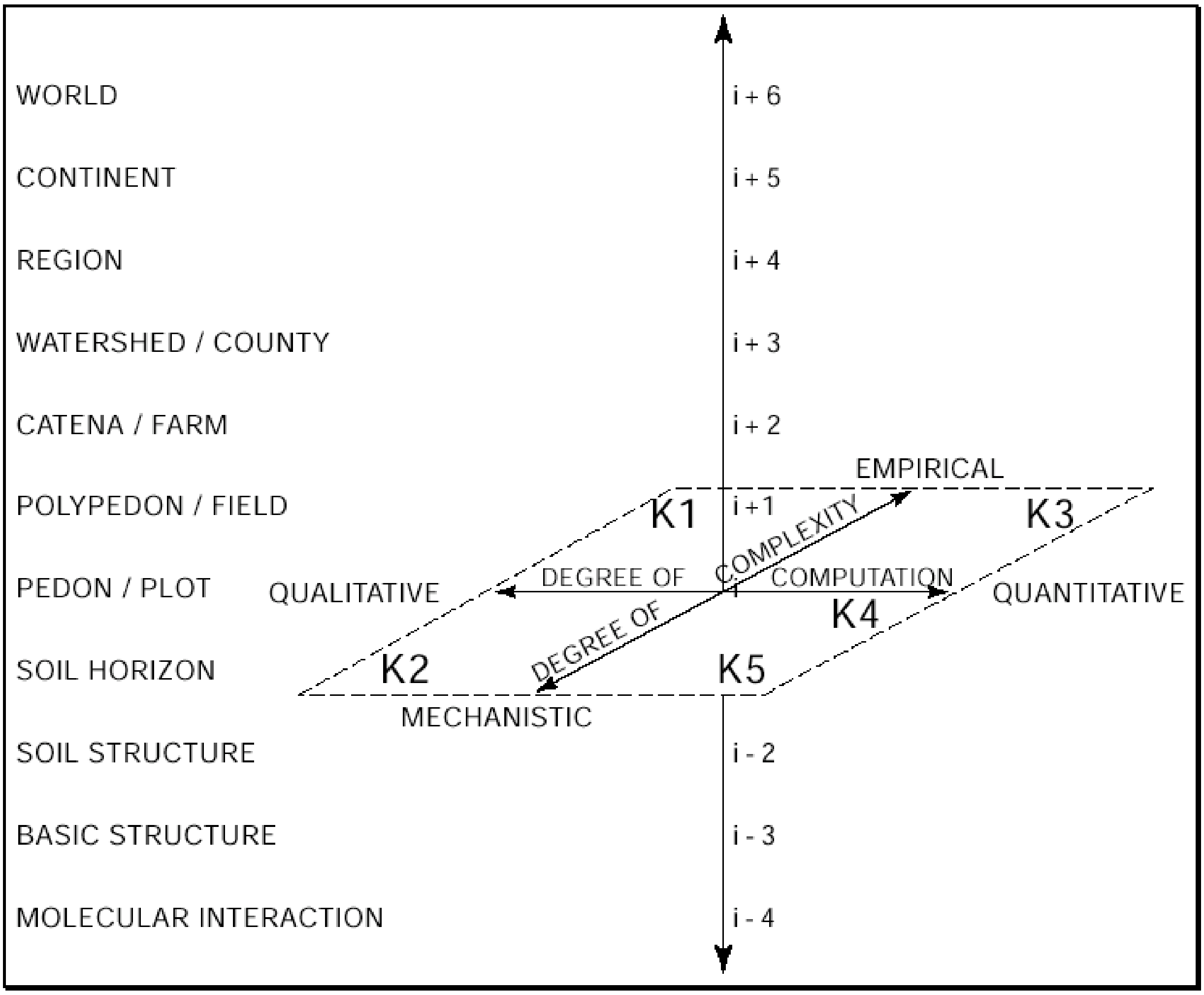

Figure 2: A decision tree for the Land Quality 'Risk of P-fixation', after Sánchez P. A., Couto W., & Buol S. W. (1982). The fertility capability soil classification system: interpretation, applicability and modification. *Geoderma,* **27**(4), 283-309.

» *ratio of free Fe_2O_3 to clay in the topsoil*
...... [< 0.15] .— LOW P-FIXATION
...... [≥ 0.15] .» *percentage of clay in the topsoil*
....[<35%] .— LOW P-FIXATION
....[≥35%] .— HIGH P-FIXATION
....[Unknown] ..» *hue of the topsoil matrix*
....[R, 2.5YR, 5YR, 7.5YR] » *topsoil structure*
....[granular] — HIGH P-FIXATION
....[other] — LOW P-FIXATION
....[10YR, Y, G, B] — LOW P-FIXATION
...... [Unknown] ..» *hue of the topsoil matrix*
....[R, 2.5YR, 5YR, 7.5YR] » *topsoil structure*
....[granular] —HIGH P-FIXATION
....[other]... — LOW P-FIXATION
....[10YR, Y, G, B] — LOW P-FIXATION

Diagnostic LCs are introduced by '»' and *italicised*.
Values of the diagnostic LCs are [boxed].
The level in the tree is indicated by the leader characters, '.....'.
Results (severity levels) are introduced by '—' and written in SMALL CAPITALS.