A new process to identify the weeds of La Réunion Island: the AdvenRun system

Thomas Le Bourgeois[1], Eric Jeuffrault[2], Pierre Grard[3] and Alain Carrara[4]
[1] Cirad, UMR PVBMT Cirad-Université de la Réunion, 3P route ligne Paradis, 97410 Saint-Pierre, Réunion
[2] Service de la Protection des Végétaux, DRAF, 2 route ligne Paradis, 97410 St Pierre, La Réunion
[3] Institut Français de Pondicherry, 11, St Louis Street, PO Box 33, 605001 Pondicherry, India
[4] Cirad-ca, Avenue Agropolis, TA74/09 34398 Montpellier Cedex 5, France

Summary Weeds are an important constraint for Réunion agriculture. Control options vary depending on the weed species present. The difficulty encountered by non-botanists when identifying weeds using classical tools such as floras or handbooks (too technical, not effective for seedlings or incomplete specimens, not sufficiently precise) led us to develop new computer aided identification tools. The software AdvenRun V.1.0 allows anybody to easily identify the major weeds of La Réunion Island. It uses a graphical system that identifies the plants step by step using images. The user has free choice of characters to describe the weed using graphical modalities. Species are listed according to the probability of each being correct. Each species is completely described with numerous colour illustrations, botanical drawings and descriptive text. Species descriptive files can also be accessed from a website where the information (e.g. distribution and control methods) is regularly updated.
Keywords La Réunion Island, weeds, identification, computer, graphical system, matching.

INTRODUCTION

In order to save non-botanists encountering difficulties when identifying weeds using standard flora, we developed a specific computer tool for species recognition and information dissemination on weeds.

How do I control weeds in my field? This recurrent question of Réunion farmers, invokes two further questions. Which weeds are there in my fields? And how can I identify them easily and precisely?

Each time a farmer has to talk about weed management with anybody else he has to describe the weed community and the main species.

Weed species names are necessary to communicate efficiently and to look for precise information. Yet, it is not so easy to recognise all the species and to give the correct species name.

La Réunion Island, a French island in the Indian Ocean, is characterised by very diverse ecological situations. Since the first human colonization, 300 years ago, more than 2000 exotic plant species have been introduced from many parts of the world (Africa, Latin America, Asia, Europe). About 600 species are naturalized, whereas the indigenous vegetation includes around 700 species (Lavergne 1999). Thus, La Réunion's flora, and especially the weed flora, is very diverse and difficult to fully appreciate.

Weed identification is difficult. Common names are sometimes ambiguous while Latin names are not widely known. Using classical floras is not easy for non-botanist people and it is not possible to identify plants at an early stage of growth. Manuals are sometimes not precise enough.

That is the reason why, in the 1990s, we decided to develop new computer tools to help people identifying weeds at any stage of growth, without any need of botanical knowledge. We used the identification by matching process. It proceeds with a step-by-step dynamical construction of a composite picture of the plant combined with a calculation of similarity coefficients. This system has been used to develop a specific tool, AdvenRun, for La Réunion's main weed species.

CLASSICAL IDENTIFICATION TOOLS AND THEIR LIMITS

Oral learning Oral learning is very frequent, using common names. But many species have several common names e.g. *Achyranthes aspera* L. with six common names in La Réunion (Herbe-d'Eugène, Herbe-des-jeunes, Queue-de-rat, Herbe-d'Inde, La Zinde, Herbe-zinde). The same common name can be used for several species, e.g. Colle-colle used for *Desmodium incanum* DC. and *Siegesbeckia orientalis* L. (Le Bourgeois *et al.* 1999).

Classical floras Classical floras show several drawbacks in their use for weed identification.

- The use of natural classification based on the hierarchy of characters, mainly sexual characters (Hutchinson *et al.* 1972). Thus, it is not possible to identify a plant at a vegetative stage.
- An identification process based on a dichotomous key system. This process is statistically the most rapid and reliable means of distinguishing many species with any certainty (Shannon 1948). But two difficulties can occur with this system. (1) It

is not possible to answer a question concerning a character not present or not clearly visible on the sample. (2) If the user makes a mistake answering a question, he or she will never get the right taxon.

- The use of technical terms not understood by non-specialists.

Manuals Some field manuals concern La Réunion's weeds (Centre d'Etude de Recherche et de Formation 1977, Le Bourgeois *et al.* 1999). These manuals vary in their precision and efficacy for weed identification or in plant descriptions and illustrations. Identification can be done using a graphical key based on vegetative characters (Le Bourgeois *et al.* 1999) or just taking a look at pictures until finding one similar to the collected sample. The reliability of confirmation depends on the quality of the description.

BACKGROUND OF COMPUTER TOOLS FOR WEED RECOGNITION

In the 1970s the idea emerged of developing computer-based tools for plant or weed identification.

The arrival of multimedia products on CD-ROMs or web sites opened opportunities to build new tools for plant identification.

Several avenues had been followed.

- Classical identification keys using text or graphics and dichotomous or multichotomous systems, i.e. Malherb software for French weeds using a graphical multichotomous key system have been developed (Lonchamp *et al.* 1991).
- In the 1980s the Delta computer convention was used to standardise plant descriptions (Dallwitz 1980). In this way, data are managed in matrix and identification keys can be automatically generated (Pankhurst 1970, 1988). Much software has been developed following this concept, notably using three different processes: Intkey (Dallwitz 1993), Lucid (Centre for Biological and Information Technology 2002, 2003, Krings 2003) or Polyclave (University of Toronto Department of Botany 2003).
- Expert system tries to simulate human reasoning. It uses an event database, rules applicable to events and an inference motor that activates rules according to events. This has been used for weeds in the software Sitrema (Sistema para el tratamiento Integral de Malezas) in Argentina (Casali *et al.* 1998).
- Neuro-mimetical networks have also been used to identify plants by leaf-shape recognition (Angel 1995).

THE ADVENRUN IDENTIFICATION SYSTEM

The process used in AdvenRun V.1.0 (Le Bourgeois *et al.* 2000) is the identification by matching. It consists of a step-by-step dynamical construction of a composite picture of the plant. This process can be easily explained following the identification of the weed *Argemone mexicana* L. At the beginning, a theoretical composite picture of a plant proposes all the characters that can be described e.g. root system, stem, hairiness, leaf, different parts of the leaf, general aspect (Figure 1). From this picture, clicking on one part of the plant, the user is free to describe the traits he or she wants, and especially the traits that are thought to be specific to the species.

Description of traits is done graphically, without technical terms, by comparing the trait shape of the sample and graphical modalities proposed, in our case the shape of the insertion of the leaf on the stem (Figure 2).

The confirmed trait is then updated on the composite picture of the plant. At each step of the description, a similarity coefficient is calculated for all the species. In our case, at the moment, there are three species with 100% of similarity. Then other traits can be described in any order (Figure 3).

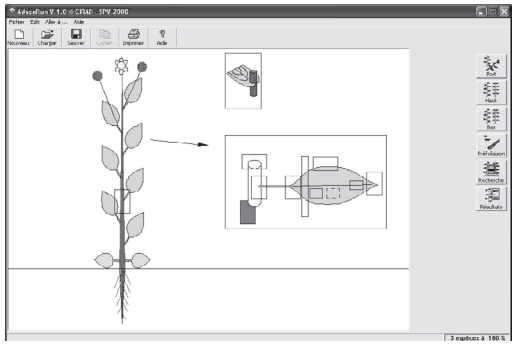

Figure 1. Theoretical composite picture.

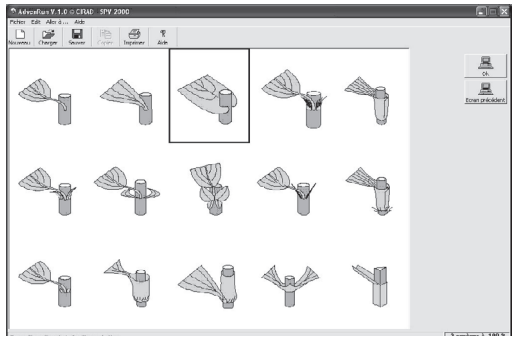

Figure 2. Graphical choice of the trait leaf insertion.

A trait that has already been described can be modified at any time without modifying any other traits.

Clicking the result button, species are listed by decreasing order of similarity and the descriptive file of the first one is proposed (Figure 4).

When the user cannot choose a new trait to describe, he can ask the computer (Research button) to propose the most relevant character to distinguish between species in the group with the highest coefficient of similarity, using the research button.

If a species has less than 100% of similarity, the composite picture indicates (red arrows) which are the characters wrongly described.

For every species and at any time, the user can get the descriptive file. In this way, it is sometimes quicker to compare the files of the four or five species at the top of the list rather to continue description of characters.

This process is highly flexible and easy to use. It is tolerant of a lack of information (incomplete sample), intraspecific variation and even observation errors.

Descriptive files of species are generated in html format. These files are accessible directly from the CD-ROM or can be available through Internet or Intranet Website with any type of browser.

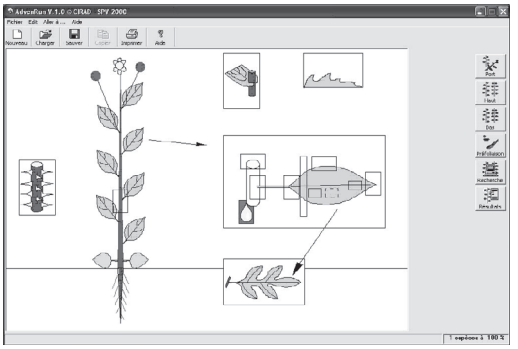

Figure 3. Composite picture of *A. mexicana.*

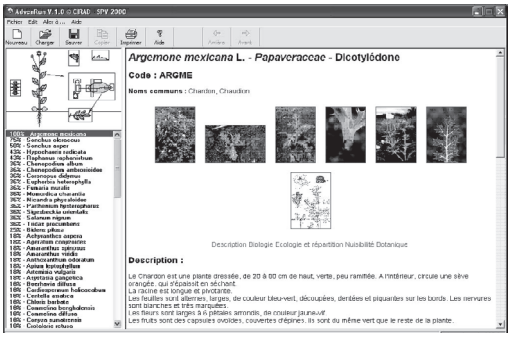

Figure 4. Species list and descriptive file.

Each species is named by its correct Latin name, main synonyms and main common names.

It is illustrated by three to eight colour slides for seedling, adult, details and full botanical drawings The species description contains a first brief, easy to understand description and then a complete botanical description. All parts of the plant are described from the seedling to the adult plant (roots, stem, leaves, flowers, fruits and seeds). In this description, all technical terms are highlighted and with a click of the mouse, an hypertext illustrated definition is accessible. Information on ecology, biology and agronomic effect on local crops is mentioned.

All documents (file, photographs, drawings) are printable.

These files are accessible from the identification system or directly from the list of plants according to family names, Latin names or common names.

Confirmation of the identification has to be done by the user comparing photos, drawings and botanical description with the sample.

In case of identification of a species not available in the software, the identification process can access a quite similar species. But, looking at the photos and drawings, and reading the botanical description the user is able to confirm if the identification is correct or not. For some similar species, a special paragraph informs that there is a species with which it may be confused and that can be distinguished by specific characters.

AdvenRun V.1.0 is a software package to identify and obtain information on about 100 of the most important weeds of La Réunion. They are mainly weeds of cabbage, sugarcane and pasture crops.

DISCUSSION

This computer tool will not be compared with others seen before because the efficiency of this kind of tool depends on the combination of three major elements:

(i) the efficacy of the identification system;
(ii) the quality of species information; and
(iii) the ergonomics of the tool.

It is then difficult to judge a tool as a whole because the user, according to his or her botanical knowledge or to his or her use of the software, could emphasise one element or the other. Furthermore, Dallwitz tried to compare seven software packages (Dallwitz 2000) but was criticised by Thiele, who showed how difficult it is to objectively compare software when the use and the specificity of each one is not completely managed (Thiele 2000).

Software has the advantage of being distributed on CD-ROMs, being accessible on the Internet or both. The accessibility to the species description through Internet is really important. It allows regular updating of the information that is essential for weed management.

One constraint of software is the need for a computer and/or Internet. It is not possible to bring a computer into the field, even a laptop, in tropical areas. At the moment, we are working on the adaptation of software such as AdvenRun to PDA (Personal Digital Assistant) which represents a new opportunity.

CONCLUSION

Weed identification is the first step in a weed control process. This step is now facilitated by new specific tools that are easy to use by farmers, extension people, students, teachers or young scientists, and not exclusive to botanists or weed scientists. AdvenRun V.1.0 at the moment deals with 100 weed species and is only in French. The next version will include about 200 species from the Western Indian Ocean Islands and will be bilingual (French/English).

Long-term weed management needs a good knowledge of weed growth processes according to ecological and agronomical components. This information and weed control methods can now be updated regularly on a Website and are accessible through Internet. This combination of weed identification and weed information dissemination will allow practitioners to adapt their practices to provide long term integrated weed management.

ACKNOWLEDGMENTS

The authors wish to acknowledge Eric Blanchard, Hélène Lefèvre, Nathalie Pas, Emmanuelle Picard, Gilles Wuster from SPV and Pascal Marnotte from Cirad for their helpful contribution. This work was financially supported by FIDOM.

REFERENCES

Angel, G. (1995). Aide à l'identification d'arbres de Guyane française par des techniques d'analyse d'images et d'intelligence artificielle – Application à la reconnaissance de forme de feuilles. Thèse de doctorat (Montpellier II, Montpellier, France).

Casali, A., Corti, R., D'Agostino, E., Dip, L., Faccini, D., Leguizamon, E., Pluss, J.J. (1998). Sistema interactivo para el tratamiento integral de malezas. I. Identificacion de malezas. *Pesquisa Agropecuaria Brasileira* 33, 809-21.

Centre for Biological and Information Technology. (2002). 'Suburban and environmental weeds'. (University of Queensland, Brisbane).

Centre for Biological and Information Technology. (2003). 'Declared plants of Australia'. (University of Queensland, Brisbane).

Centre d'Etude de Recherche et de Formation. (1977). Les principales adventices de la canne à La Réunion. (CERF, Saint Denis, La Réunion, France).

Dallwitz, M.J. (1980). A general system for coding taxonomic descriptions. *Taxon* 29, 41-6.

Dallwitz, M.J. (1993). DELTA and INTKEY. *In* 'Advances in computer methods for systematic biology: artificial intelligence, databases, computer vision', ed. R. Fortuner, pp. 287-96. (Johns Hopkins University Press, Baltimore, London, UK).

Dallwitz, M.J. (2000). A comparison of interactive identification programs. http://biodiversity.uno.edu/delta/www/comparison.htm.

Hutchinson, J., Dalziel, J.M., Keay, R.W.J., Hepper, F.N. (1972). 'Flora of west tropical Africa'. (The Whitefriars Press, London and Tonbridge, UK).

Krings, A. (2003). 'Common grasses of Carolinas. Applied taxonomic solutions'. (Raleigh, USA).

Lavergne, C. (1999). The invasive woody weed *Ligustrum robustum* subsp. *walkeri* threatens native forest on La Réunion. *Biological Invasions* 1, 377-92.

Le Bourgeois, T., Jeuffrault, E., Fabrigoule, S. (1999). AdvenRun – Principales mauvaises herbes de La Réunion, Description-Herbicides. (Cirad/SPV, Saint André, La Réunion, France).

Le Bourgeois, T., Jeuffrault, E., Grard, P. and Carrara, A. (2000). AdvenRun V.1.0. – Principales mauvaises herbes de La Réunion. (Cirad – SPV, Montpellier, France).

Lonchamp, J.P., Barralis, G., Gasquez, J., Jauzein, P., Kerguelen, M., Leclerch, J., Maillet, J. (1991). Malherb, logiciels de reconnaissance des mauvaises herbes des cultures: approche botanique. *Weed Research* 31, 237-45.

Pankhurst, R.J. (1970). A computer program for generating diagnostic keys. *Computer Journal* 12, 145-51.

Pankhurst, R.J. (1988). An interactive program for the construction of identification keys. *Taxon* 37, 747-55.

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal* 27, 379-423 and 623-56.

Thiele, K. (2000). A critique of Dallwitz's 'A comparison of interactive identification programs'. http://biodiversity.uno.edu/delta/www/thiele.htm.

University of Toronto Department of Botany (2003). Welcome to Pollyclave. University of Toronto Department of Botany. http://prod.library.utoronto.ca/polyclave/index.html.