

CROPS, WEATHER AND SATELLITES: INTERFACING IN THE JUNGLE^α

R. Gommes

Agrometeorology Group
Environmental Information Management Service
FAO

Via delle Terme di Caracalla, 00100 Rome, Italy
phone: +39 6 52254121, fax: +39 6 52255731
email: rene.gommes@fao.org

Abstract

FAO has been involved in crop monitoring and forecasting for food security since 1978, when the Tanzanian Early Warning System was established. This was to be followed by about 40 other national and regional food security systems.

Methodological developments originally took place independently at the national level leading to a variety of software applications, mainly in the field of agrometeorology. In 1988, FAO started disseminating remotely sensed products (images and software) at the national level.

Food security programmes rely on other data types as well (markets, prices, food-stocks, agricultural statistics), and a number of other partners are involved in Early Warning activities (NGOs, etc.). Standardisation is essential if applications are to be compatible, if they are to evolve, if people are to be trained in their use and if analyses are to yield comparable results..

Given the number of actors, it is virtually impossible to reach an agreement on software standards. Therefore, it has been the philosophy of FAO to try and standardise data files instead. If, in addition, software can be designed as a succession of self-contained modules, applications can almost be "composed" by assembling existing building blocks (i.e. modules).

It is suggested that the benefits of the approach outlined above by far outweigh the costs, if only in terms of the credibility of the various partners.

INTRODUCTION

The title of this paper refers to a situation observed in relation with the FAO food security programs. Such programs are usually known as “Early Warning Systems” (EWS), but they come under a variety of names (¹).

Early Warning Systems monitor all the aspects of food availability, and the factors which affect it. They now typically include three elements: (i) an economic component dealing with food marketing, storage, imports and exports, etc; (ii) a component monitoring the national food production (agrometeorology, remote sensing and agricultural statistics) and (iii) a household food security and nutrition component which assesses food availability at the level of final consumers, i.e. individual households.

In Early Warning Systems, agrometeorology and remote sensing play an important role in crop monitoring (qualitative) and forecasting (quantitative). FAO has developed in parallel a series of methods and software tools to improve analyses and solve common problems encountered. A detailed account of the FAO software families is given in Gomme (1995).

The word “jungle” in the title refers to several rather different situations. A “jungle” of national practices results from the traditions of national meteorological and agrometeorological services, which all evolved independently, sometimes in opposite directions. Next, a “jungle” of data categories (²) from the broad spectrum of weather to food prices and from food production statistics to satellite indices, which are routinely used in EWS. This is a situation which cannot be avoided, but tools must be designed to bring all the data to a common denominator for their final analysis. Finally, a “jungle” of tools and formats, more specifically linked with the processing. It was also suggested at the COST meeting in Budapest that there might be a “jungle” of attempts to standardise. This is a more difficult issue, which FAO tried to contribute solving by organising a meeting on the *Standardization and harmonization of databases and software for agroclimatic applications* (FAO, 1995), but more work is obviously required on the subject.

THE JUNGLE OF NATIONAL PRACTICES

Professional stations tend to strictly follow WMO recommendations as regards observation times, standards and reporting procedures.

Agrometeorological services, which are usually the direct partner of the national EWS, tend to rely more on volunteer observers; most of them have adopted local data collection standards which may vary significantly from the practice recommended by WMO and their neighbouring countries.

¹ For instance "Food Early Warning and Information System (EWFIS)", "Crop Monitoring and Early Warning System(CMEWS)", "Crop Forecasting and Monitoring System (CFMS)"...

² Categories refers to data of different natures, for instance meteorological data and crop statistics.

For instance³), some countries collect rainfall data by pentads (5 day periods) starting on 1 January and carry the system throughout the year (thus 1-5 Jan., 6-10 Jan., 11-15 Jan., 16-20 Jan., 21-26 Jan., 26-30 Jan., 31 Jan.- 4 Feb., 5-9 Feb., etc); others re-start the system on the first of each month, thus inserting a 6 day pentad every second month (1-5 Jan., 6-10 Jan., 11-15 Jan., 16-20 Jan., 21-26 Jan., 26-31 Jan., 1-5 Feb., 6-10 Feb...).

Others still use weeks starting on 1 January (1-7 Jan., 8-14 Jan., 15-21 Jan., 22-26 Jan., 27 Jan. 2 Feb., 3-9 Feb., which has the advantage that weeks cover the same dates every year. The disadvantage for the routine of agrometeorological services is that the weeks end on a different day every week, causing delays in transmission when the week end over the local week-end (Saturday-Sunday, Thursday-Friday or Friday-Saturday). A second alternative with weeks is to have them cover the days from Thursday to Wednesday (or Tuesday to Monday), which is perfect for the routine, but has weeks end on a different date every year.

Of course, the introduction, by the Commission of Agricultural Meteorology (CAgM) of the dekad, the standard 10-day period, has considerably clarified the situation, as the dekad now tends to be adopted by many agrometeorological services. However, the new difficulty which arises is that the dekad is probably a bit long for some impact assessments (dry spell up to 19 days can go unnoticed).

The dekad is also frequently adopted as a reporting period only, i.e. daily rainfall (or temperature, etc.) is being reported, but at ten day intervals. This is not far from an ideal compromise.

It remains that application software, say a crop water balance program, cannot possibly be designed to cope with all the existing national situations. It is thus likely that, to some extent, countries will have to adapt to tools.

THE JUNGLE OF DATA CATEGORIES

Data categories constitute a complex issue. In the area of food security, they are usually drawn from different technical fields, such as agrometeorology, pedology, remote sensing, marketing, *etc.* Table I lists some of the data together with their major sources and characteristics. The institutions responsible for measuring and recording the data vary greatly, but this is usually a minor complication in comparison with the profound differences in the nature of the data themselves, particularly their sizes in terms of geographic area covered by the samples and the sampling frequency. Ideally, the sampling frequency should be inversely related to the stability (in space and time) of the variables. This is not always the case.

Three more points will be mentioned in this context: the data “types” inside each category, the adequacy of the data for Early Warning purposes, and some of the problems linked with missing data.

³ This discussion does not attempt to present the systems adopted to deal with leap years.

Technical field <i>(Data Categories)</i>	Ministry <i>(Source)</i>	Data type <i>(Variables)</i>	Geographic Unit	Sample size <i>(area)</i>	Sampling Frequency	Sample size <i>(time)</i>
Agro-meteorology	Transport, Environment, Natural resources, Defense, Agriculture...	Weather data	Station	1+	Day	1
		Phenology			Week	10
		Crop condition	Region	1000	Week	10
		Phenology			Month	50
		Phenology			Week	10
Soils	Agriculture, Nat. resources	Soil WHC	Station Polygon	1 10-100	10-20 years	5000
Remote Sensing	Transport, Defense, Research	NDVI (Crop condition, phenology) CCD (Rainfall)	Pixel 7 x 7 km	10	Hour, then aggregated to week	0.1-1
Marketing	Commerce, Agriculture	Price data	Market	5+	Bi-monthly	20
		Grain stocks	Warehouse	5+	Day	1
		Grain	Village	5	Year	500
Agricultural Statistics	Agriculture	Cropped area Yield Production (Crops and livestock)	Admin. unit, Agric. dev. district, or "polygon 1"	100	3- to 12-monthly	200-500
Demography	Planning Central statistics	Population	Administrative Unit, or "polygon 2"	100	5 years	2000
Nutrition	Health	Children's weight, other indicators	Dispensary	5+	Weekly to monthly	10-50

Table I: Main data used in Early Warning Systems, by categories, together with their institutional sources. The last columns indicate the area and duration for which the sampled data are considered representative.

The “types” of data (or “variables”) belonging to the same category are relatively well defined, for instance rainfall and temperature in the “weather data” category. However, countries often adopt different “styles” for the same variables, for instance different units, which cannot always be converted. For example, grain prices can be expressed in price per weight (e.g. 100 crowns per kg) or in price per volume (e.g. 1000 shillings per bag, where bags may have different sizes). Things tend to be easier with the physical environment where - for instance, miles/day are easily converted into m/s - but problems frequently arise with some remotely sensed variables (e.g. cloud data) which cannot immediately be related to ground observations for a number of reasons linked with different types of sensors and dissimilar sample sizes - both spatially and temporally -.

The second point (adequacy of data for Early Warning purposes) has to be looked at from two different levels: the design of the data collection systems and the types of data collected. First of all, data collection schemes have almost never been designed to serve the specific purpose of food security. A good and classical example is the location of many weather stations, which rarely follow the pattern of agro-ecological zones and major areas of food production. Examples could be multiplied.

Regarding the “styles” of data, two examples are given: (i) farm-gate crop prices and (ii) crop phenology. Farm-gate prices are prices actually paid by rural people, i.e. the majority of people in developing countries, for their food. They are more relevant than prices collected in central and urban market places. Crop phenology observed at agrometeorological stations often gives a picture that is not representative for the farmers’ fields, due to different practices - planting dates, inputs - and different varieties. It is most striking when agrometeorological stations in African highlands report the phenology of tomatoes, egg plant and cabbage, when the whole area depends on maize for their food. Except for scientific purposes, it is not very meaningful either to report detailed phenological phases (⁴) (1st, 2nd, 3rd leave...) which are of very little relevance for crop monitoring and forecasting.

The third point in this section - missing data - covers a very complex subject which would deserve a separate treatment. To start with, there are different categories of missing data. Many data termed “missing” are, in fact, “not relevant”. For instance, if no maize is cultivated in district A, both cultivated area and production are 0 (nil), but yield (production/area) is not relevant.

The example above also touches on the uneasy relations between 0 (nil) and “missing”, which may even creep into the physical data: “provisional” rainfall amounts are often used in water balance calculations, when dekads are used and the last day(s) have not been received yet. This leads to no serious consequence if amounts exceed soil storage capacity, and when the data bases are managed very carefully. This is, however, not always the case.

⁴ Interestingly, “harvest” is not considered a phenological phase, although the information is essential for crop monitoring and by no means always coincides with full maturity. Due to a number of factors, there may be months between full maturity and harvest.

Missing data sometimes carry at least as much information as actual ones. One of the best examples can be taken from the field of marketing: prices sometimes keep rising for weeks, due to short supply, until the commodity eventually disappears from the market. The data item information, in that case, is technically missing, but it remains very relevant and paradoxically carries a lot of useful information. A parallel example can be taken in the field of demography: human populations temporarily or permanently migrate out of the villages in the case of very serious food crises. Is the population now “nil”, missing or “not relevant” ? This example clearly shows the importance of the spatial and temporal aspects (⁵).

Finally, this section has to mention data quality, in that there exists, for some data types, a continuum between “good” data and missing data, with data of decreasing quality in between. The partial rainfall total mentioned above is one example. Another is satellite based vegetation indices which are liable to being affected by “sub-pixel” cloud contamination. Only in obvious cases will this be detected, leading to data being actually of unknown quality, from “probably good” to “probably bad”.

It is clear that Geographic Information Systems play a crucial role in providing a common denominator between the categories of data mentioned. They do not, however, offer solutions to most of the problems listed. Converting data to the common denominator will, in fact, require tools, which takes us to the next “jungle”.

THE JUNGLE OF TOOLS (⁶) AND FORMATS

What is meant here is not the large array of tools needed in Early Warning, and specially agrometeorological practice, for instance the general tools needed to estimate missing data, model crops, compute soil moisture, area averaging and gridding, calibrate models, visualise and extract data from satellite images, *etc.* Such tools are necessary, and it is a positive factor that many different tools should be available to solve specific problems.

The major obstacle is that the tools can exchange data only with difficulty in most cases, including in the case of the above mentioned common denominator, i.e. GI Systems. This is a typical example of a new solution fraught with a number of equally new problems.

Operational work thus requires efforts to have the tools work together, mainly because most of them come with specific, and often proprietary, data formats (⁷). Converting file formats has

⁵ The population of district A migrated into another area B, where the population increased, even if the newcomers interacted only little with the local population. The population of B remains a well behaved variable. In A, however, the population may become “not relevant” rather than 0 (nil) if the out-migration is permanent.

⁶ This refers specifically to computer tools, i.e. software. Hardware problems, such as those arising magnetic media used for recording data in automatic weather stations, are not covered.

⁷ There is some ambiguity regarding this term. The following elements contribute to defining a data file format: the type of information (with qualifiers, such as units and code used for “missing” data), the way in which the information is arranged, the file naming conventions. It is clear that properly designed standard files should be self-describing.

become a new branch of data processing, and the field of Early Warning Systems is certainly no exception.

While more or less efficient conversion programs abound for general purpose software (word processing, spreadsheets, graphic formats), an even greater variety of conversion programs was developed for specialised areas. Unfortunately, and this constitutes almost a rule, such software is developed locally and on an *ad hoc* basis, whenever they are required, sometimes at great expense, but always at the cost of the time that could otherwise have been dedicated to analyses. Needless to say, the same conversion software is developed independently many times.

Obviously, the solution is the adoption of standard data formats. Some standards have started emerging, e.g. CLICOM for climatological data base management. Unfortunately, many standards are short-lived due to the fast development of hardware and software, and the fashions that affect informatics.

One of the difficulties is thus to design data standards which have a reasonably good chance to last.

SOME CONCLUSIONS AND SUGGESTIONS

There is a need to clear the above-mentioned “jungles”, and this will be possible only through good will and collaboration, starting with data providers and data users. The new partners in this discussion, the software developers - among which GIS expert occupy a central niche-, stand between data providers and users. It is suggested that they also hold the key ⁽⁸⁾ to reducing the enormous waste of human and financial resources which accompany the lack of standardisation.

The central idea behind this paper is that the main problem is the standardisation of data files. If this can be achieved, than the large variety of application software, mainly in agrometeorology and remote sensing, will be compatible, almost as a by-product of data standardisation.

Below are some of the feature that describe “standard data files”: (i) they are self descriptive (name, contents, structure, format), (ii) they are **not** user-specific, (iii) they are not program or application specific, (iv) they are not location specific (cropping seasons) and they are language tolerant and, finally (v) they are not proprietary.

It is also suggested that software should adopt a modular approach, i.e. it should ideally be regarded as the succession of interlocking building blocks. Needless to say, this must be kept in mind at the development stage. The advantages of modular software include (i) reduced development overhead, (ii) reduced training overhead, (iii) ease of maintenance (even by different people!), (iv) ease of upgrading and (v) robust and sustainable systems.

⁸ Among others because they are largely responsible for the present situation.

Standard data files and modular application software, together with the complete separation of data management and data analysis are proposed as the three keys to solving most of the problems listed above.

The final question thus remains: who has the necessary competence, authority and resources to take the lead?

REFERENCES (⁹)

FAO, 1995: Standardization and Harmonization of databases and software for agro-climatic applications. Proc. of an expert meeting held in FAO, Rome, 28 Nov. - 3 Dec. 1993. *FAO Agrometeorology series working paper* N. 13, 313 pages.

Gommes, 1995. A note on FAO Early Warning Software. *In*: FAO, 1995: 111-126.

^α Gommes, R., 1996. Crops, weather and satellites: interfacing in the jungle. Proc. of COST 77 workshop on **The Use of Remote Sensing Techniques in Agricultural Meteorology Practice**. Budapest, 19-20 Sept. 1995. Pp. 89-96 of EUR 16924 EN, EU Directorate-General Science, research and Development, Luxembourg: Office for Official Publications of the European Communities, 289 pages.

⁹ The reader is referred to FAO, 1995, for a more illustrative and comprehensive coverage of the subject.