CHAPTER 3. DATA GATHERING PROCESS

Measurements in the Chaguana river, Ecuador

Parts of this chapter were published in

Matamoros D., Bonini J., Guzman E., Ramirez G., and Vanrolleghem P.A. (2002). Design and implementation of a measuring campaign to model pesticide impacts in an Ecuadorian watershed. Water Science and Technology, 46(9), 37-45.

Matamoros D., Van Biesen L. and Vanrolleghem P.A. (Submitted). Development of a Geographical Information System for pesticide assessment on an Ecuadorian watershed. *Submitted to:* Water Science and Technology.

3.1. INTRODUCTION

The present chapter shows the procedure followed to gather all necessary data to run the evaluated models, mainly AGNPS. When gathering information, data is not necessarily in the correct format for the model execution. Thus, this chapter also presents several processes to convert the collected data. The data structure of each model is shown in the next chapter.

As part of the research objectives, this chapter presents some guidelines for selecting, gathering and processing poor data for environmental evaluation of watersheds in developing countries based on the current experience in Ecuador. The following sections show data processing procedure followed to obtain the most reliable data to be used in the model evaluation phase.

3.2. METHODOLOGY

Before using an environmental model to determine pesticide impacts caused by banana plantations in a watershed, an important step is the collection of representative and reliable data to be used in the model. In Ecuador, like in the majority of developing countries, public and private institutions do not always maintain a good housekeeping of environmental records. In addition, available information is sometimes outdated and not easily accessible. Thus, obtaining information could be difficult to accomplish, and some extra efforts must be done to process the gathered data.

In this Ph.D. thesis, GIS procedures were used in the data evaluation process. The platform selected for GIS processing was ArcView. Most of the GIS generated data was used as input data in the model execution step. However, the majority of data representing the different phenomena are only available as scarcely and irregularly distributed data. Thus, it is necessary to convert those data to a raster format in order to consider spatial variability in the modelling work. Other data were converted to vector format[10] as an intermediate step to aggregate raster data for classification purposes. All data were grouped in two categories:

[10] Vector data in ArcView is represented by shapefiles which is a simple non-topological format for storing attributes and location of data.

- **Primary data** were collected directly from available national databases or measured in the field. These data were evaluated according to the source, year of publication and scale in order to maximise information extraction (Table 3.1). For example, edaphology data was extracted initially from printed maps (only existing soil taxonomy groups in the area); later, these data were complemented with own soil sampling campaigns (Matamoros *et al.* 2002) and lab analysis (texture, soil moisture content).

Table 3.1. Primary data obtained in the Chaguana river basin

Data	*Primary Source*	*Year*	*Scale*	*Data extraction procedure*
Topography	Printed Maps	1970	1:50000	Scan of 4 topographical sheets
Geology	Printed Maps	1970	1:250000	Scan of 1 geological sheet
Edaphology	Printed Maps	1970	1:250000	Scan of 1 edaphological sheet
Land use	Digital Data	1998	1:50000	Visualization in ArcView
Climate	Database	Depends on station	Not applicable	5 georeferenced weather stations
Soil	Field Measurement	2001	Not applicable	30 georeferenced sampling sites
Water Quality	Field Measurement	2002	Not applicable	26 georeferenced sampling sites

- **Secondary data** are more elaborated data that include spatial variability in grid format. For the current research, it was generated from primary data by using kriging interpolation, accepted equations or methodologies applied in similar situations. The majority of information was generated as raster format because the evaluated models are raster-based. The cell size for the generated raster data was 1 hectare. Multiple thematic maps were developed to extract input data for the models (Table 3.2).

Table 3.2. Thematic maps generated from primary data on the Chaguana river basin

Thematic Maps	*Geographic Feature*	*Generation procedure*	*Reference*
Elevations	Point	Digitising from scanned maps	GIS procedure
Digital Elevation Model	Grid	Interpolation procedure	GIS procedure
Slope	Grid	ArcView Avenue statement	GIS procedure
Geologic Units	Polygon	Digitising from scanned maps	GIS procedure
Taxonomic Units	Polygon	Digitising from scanned maps	GIS procedure
Weather Stations	Point	Add as an event theme in ArcView	GIS procedure
Precipitation	Grid	Interpolation procedure	GIS procedure

Thematic Maps	*Geographic Feature*	*Generation procedure*	*Reference*
Runoff Erosivity Factor	Grid	Use of RUSLE equations in Map Calculator	Renard *et al.* (1997)
Sampling Sites	Point	Add as an event theme in ArcView	GIS procedure
Clay – Silt – Sand content	Grid	Interpolation procedure	GIS procedure
USDA Soil Texture	Grid	Boolean Algebra	Benham *et al.* (2001)
Very fine sand content	Grid	Interpolation procedure	GIS procedure
Soil Moisture content	Grid	Interpolation procedure	GIS procedure
Saturation content	Grid	Use of equations in Map Calculator	Saxton *et al.* (1986)
Field capacity content	Grid	Use of equations in Map Calculator	Saxton *et al.* (1986)
Wilting point content	Grid	Use of equations in Map Calculator	Saxton *et al.* (1986)
Bulk density	Grid	Use of equations in Map Calculator	Saxton *et al.* (1986)
Saturated hydraulic conductivity	Grid	Use of equations in Map Calculator	Saxton *et al.* (1986)
Hydraulic soil group	Grid	Boolean Algebra	USDA-NRCS (1986)
Organic matter content	Grid	Interpolation procedure	GIS procedure
Soil albedo	Grid	Use of equation in Map Calculator	Baummer *et al.* (1994)
Soil Map	Polygon	Data aggregation from soil grids by using Taxonomic Map as a mask	GIS procedure

3.3. SELECTION OF THE STUDY AREA

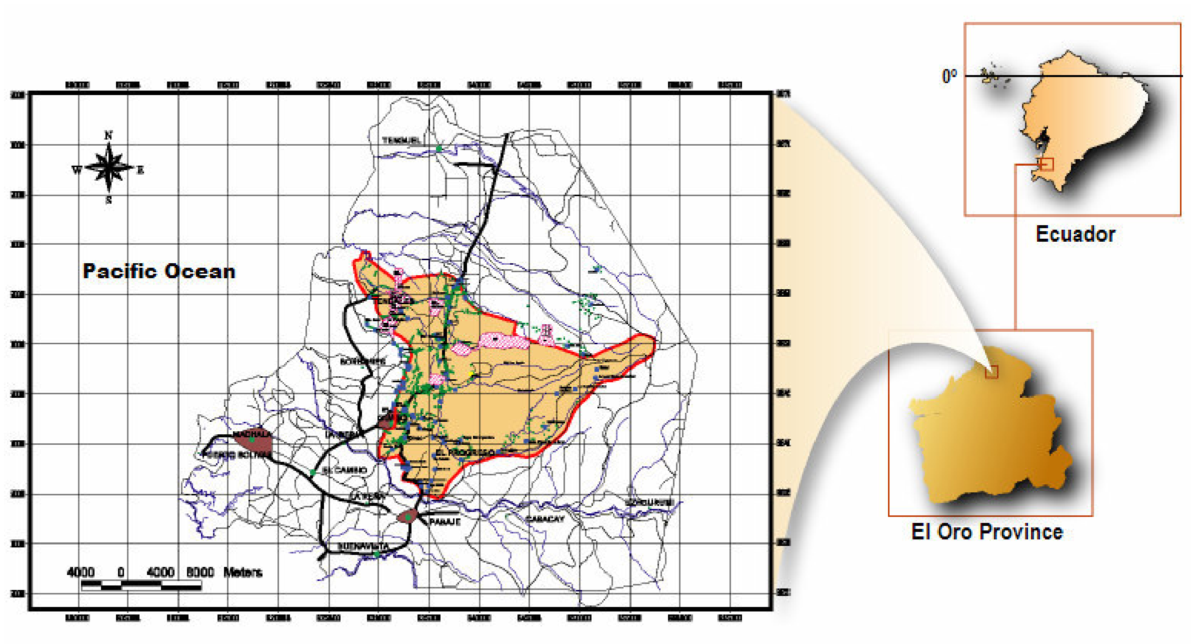
According to the Ecuadorian Central Bank, the Ecuadorian banana activity covers approximately 139000 Ha distributed over 7 coastal provinces. However, the banana activity is not always the only crop affecting a specific site. Thus, an impact assessment of pesticides only coming from banana plantations could be impossible to achieve[11]. For that reason, the current research selected a site in such a way that the main goal could be achieved by minimizing other crop interference. The criteria applied for site selection were:

- As pesticides are transported mainly through the hydrological cycle, a river basin should be the most acceptable site to perform the assessment. However, pesticide drift from adjacent basins could influence the final assessment of the selected river basin.

[11] Other crops sometimes use the same agrichemicals as banana farms do (e.g. glyphosate)

- The basin should be median or low size (less than 50000 Ha) in order to avoid multiple crop activities. The larger the basin, the bigger the chance to find more than one crop activity in the basin.
- At least some farmers should agree to collaborate with research (farm management records, pesticide usage, soil and climate records and accessibility)

After visiting some potential study sites in Ecuador, the Chaguana River Basin was selected. It is located in the El Oro Province in the South-Western part of the country (Figure 3.1). This river basin is approximately 32000 Ha, which could be considered as a hydrological system between a rural catchment and a river basin (Maidment 1996). The Chaguana system does not discharge directly to the sea, but to a bigger watershed called the Pagua River Basin. However, the sea tidal influence reaches up to 6 km upstream the Chaguana basin's outlet.



Source: Centre of Environmental Studies (CEMA) – ESPOL, 2002

Figure 3.1. Location of the Chaguana river basin

3.4. TOPOGRAPHICAL DATA

3.4.1. AVAILABLE CARTOGRAPHICAL INFORMATION

In Ecuador, the Geographical Military Institute (IGM) is the official organism to keep, publish and distribute all topographical information. Most of the available information is from 1970, and it is distributed as printed maps at 1:50000 scale. On those printed maps, elevation contours are displayed every 40 meters with scattered elevation points representing measured bench marks above sea level. Terrains between 0 and 40 meters do not show any elevation contour. Contours on printed maps were drawn based on aero-photogrammetric procedures.

Because topographical information is always used as a reference for other types of information, cartographical data were obtained from those printed maps, with geographical characteristics given in Table 3.3

Table 3.3. Cartographical characteristics used in the geo-referencing data process

Coordinate System / Projection	Universal Transverse Mercator – Zone 17 S
Datum	Provisional South American Datum 1956
Spheroid	International 1924
Horizontal and Vertical Units	Meters
Latitude of Origin	0º
Central Meridian	-81º
False Easting (meters)	500000
False Northing (meters)	10000000

The selected study site, the Chaguana River Basin, is covered by 4 printed topographical sheets: Machala, Tendales, Uzhcurrumi and Ponce Enriquez. The basin is enclosed in a rectangle whose boundary coordinates are shown in Table 3.4. The basin's centroid is located at 641000 E and 9647000 N. The maximum recorded elevation on the basin is 3267 meters and the minimum elevation is 1 meter above sea level. The banana sector in the basin is located below 60 meters and above 4 meters elevation levels.

Table 3.4. River basin boundary locations

Boundary	North	East	South	West
UTM Coordinate	9658919 N	656048 E	9634843 N	625952 E

3.4.2. GENERATION OF DIGITAL ELEVATION MODEL AND OTHER NEEDED TOPOGRAPHICAL CHARACTERISTICS FOR THE STUDY SITE

The Digital Elevation Model (DEM) is raster type data that contains spatially distributed elevation information to allow an automatic delineation of watersheds[12]. DEM for the Chaguana Basin was not available at the time of the study; thus, it was necessary to generate it based on existing topographical maps following the procedures bellow:

[12] Watershed delineation is based on the eight-direction pour point model (Jenson and Dominque 1988)

1. *Digital Conversion of Printed Maps*: Topographical information was converted into digital data by first scanning topographical sheets and then geo-referencing scanned images.
2. *Extraction of Elevation Data*: All elevation contours were digitized into a polyline coverage called *CONTOURS*. Polyline vertexes were then converted into an elevation point coverage by using the EDIT TOOL extension (Tchoukanski 2002). It was necessary to edit the elevation layer to obtain the best terrain representation, which includes manual extrapolation of additional elevation points and addition of some measured terrain levels. Scattered points were also added manually to adjust the shape of the derived river streams to the shape of existing ones[13].
3. *Cell Size Selection*: Before generating a raster image such as a DEM, it is important to define the size of the cell containing elevation data within the raster format. Dealing with watershed delineation, a rule called *thousand-million*[14] is usually applied to obtain the minimum recommended cell size for the watershed assessment (Maidment 1996). For the current watershed assessment, the regional area enclosing the basin is 750×10^6 m^2 (≈ 25 km × 30 km). Therefore, the minimum recommended cell size was 750 m^2 (≈ 27 m × 27 m). However, the cell size was defined as 10000 m^2 (100 m × 100 m) because the majority of agricultural practice management in Ecuador is performed on one-hectare basis. The selected cell size will be used in all generated raster maps.
4. *Data Interpolation*: Finally, the DEM is generated by using an interpolation method with all elevation point data. The selected method was the *Universal Kriging Interpolation Procedure* included in the KRIGING INTERPOLATOR 3.2 extension (Boeringa 2002). The main problem with interpolation is the potential creation of systematic errors in the generated surface. Therefore, the interpolation procedure should be repeated as many times as necessary together with the point layer editing process and river stream generation in order to obtain the best DEM for the study site. Figure 3.2 shows the generated digital elevation model (*FILLEDEM100* raster file) with a 100 m cell size.

[13] Further explanation on stream derivation from DEM will be explained in the hydrological data processing section 3.5.

[14] The rule states that the regional area enclosing the evaluated watershed can be divided down to one million cells without compromising too much effort in the assessment. Further watershed subdivision might not add more precision to the watershed assessment. In addition, the minimum watershed area that could be delineated in that region is obtained by multiplying one thousand times the cell area.

Once the DEM is obtained, other important topographical data for modelling purposes can be obtained by using accepted equations within Avenue statements and macro tools (Map Calculator) in ArcView.

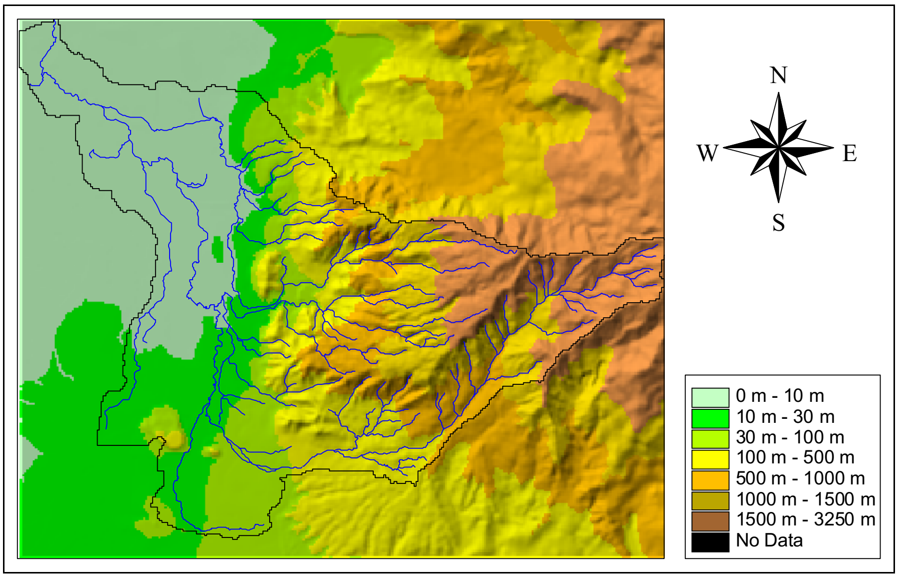


Figure 3.2. Digital Elevation Model with a 100 m cell size

<u>Terrain Slope:</u> The terrain slope angle representation (*SLOPE_PERCENT* raster file) can be generated by using AVENUE statements. The value of each cell is represented as percent rise. Figure 3.3 shows the generated slope raster image. The applied Avenue statement was

SLOPE_PERCENT = FILLEDEM100.Slope(Nil,TRUE) [3.1]

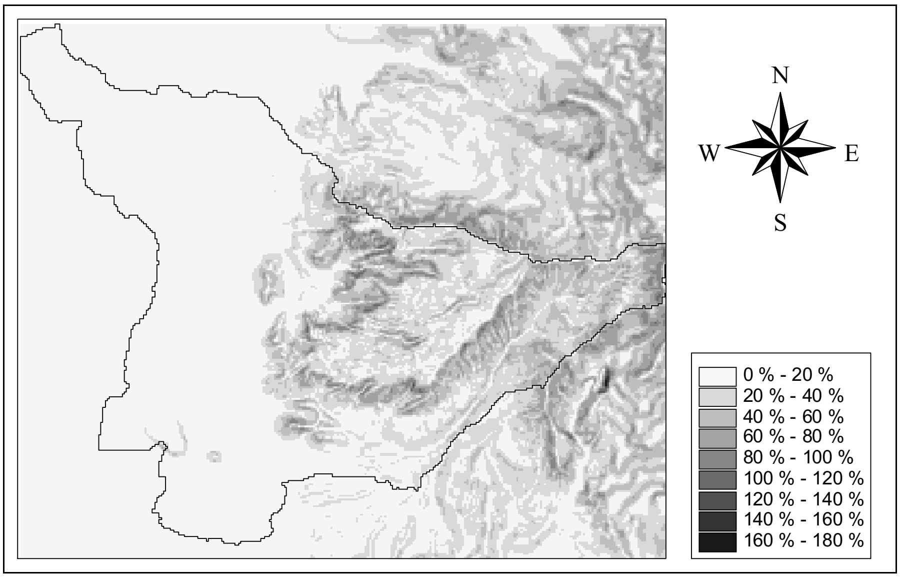


Figure 3.3. Slope percent of the Chaguana river basin

3.5. HYDROLOGICAL DATA

3.5.1. EXISTING FLOW GAUGING STATIONS

When modelling chemicals in rivers, the hydrological data is a very important piece of information to help in quantifying chemical amounts transported by the river flow. In the current research, it was possible to obtain some hydrological information (periodical water flows) in three existing gauging stations from the former National Institute of Water Resources[15]. The locations of the existing stations are shown in Figure 3.4.

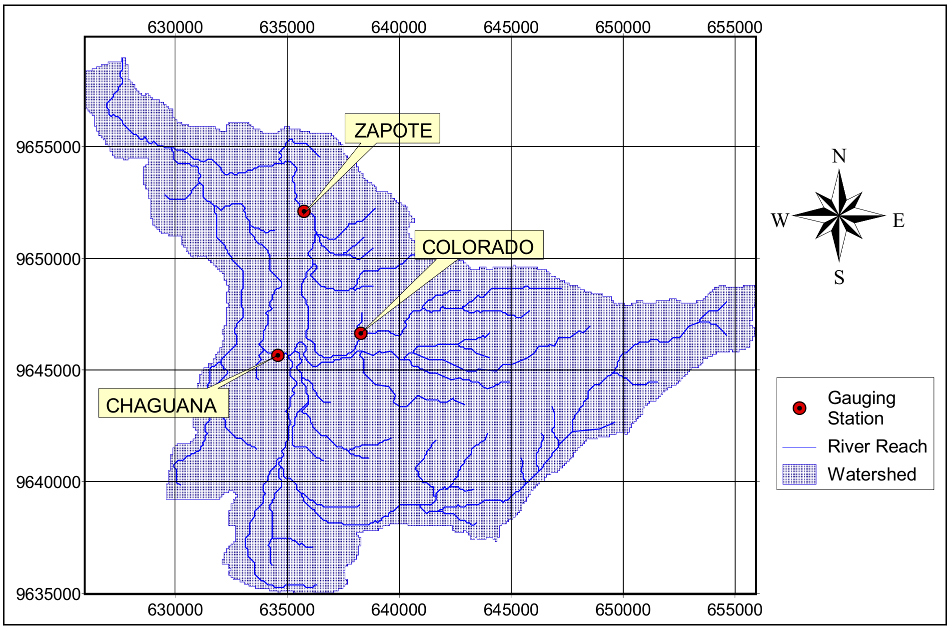


Figure 3.4. Location of existing gauging stations

However, there are some drawbacks related to these records.

- Apparently, the Water Agency is no longer measuring flow data on those stations. The supplied hydrological data only represents a 4-year non-continuous period of measurements (1978 – 1980 and 1982 – 1983).
- The data represent only average monthly flows.
- Some monthly flows are missing at all stations (probably not measured)
- There is no reliable information regarding the measurement method. However, there is a possibility that the flows have been measured by using the existing level marks attached to the bridge's piles at the gauging points (photo 3.1).

[15] INERHI (Spanish acronym) was changed by law in 1994 to the National Council of Water Resources (CNRH)

Photo 3.1. Level marker attached to a bridge's pile

For that reason, the available data was statistically processed to obtain average flow values to be used only in the hydrological calibration of the model. Table 3.5 shows median values of recorded flows within the historical period. The reason of using the median is explained later in the evaluation of meteorological records. From the table, April is the month with the highest flows recorded in the basin. This is strongly related with the rainy season in Ecuador lasting from January to May (see Section 3.6).

Table 3.5. Median flow measurements on existing gauging stations (1979 – 1982)

Gauging Station	*Drainage Area (Ha)*	*Measured monthly flows (m^3/s)*											
		JAN	*FEB*	*MAR*	*APR*	*MAY*	*JUN*	*JUL*	*AUG*	*SEP*	*OCT*	*NOV*	*DEC*
Zapote	5772	0.704	0.491	1.125	1.161	0.756	0.483	0.596	0.451	0.363	0.297	0.690	0.669
Colorado	2422	0.350	0.551	0.731	1.813	0.859	0.478	0.513	0.306	0.252	0.284	0.209	0.268
Chaguana	16199	0.762	1.440	1.646	4.453	1.621	1.121	0.619	0.571	0.346	0.481	3.470	1.173

3.5.2. FIELD FLOW AND CROSS SECTION MEASUREMENTS DONE IN THE RESEARCH

Due to limitations in the flow data, a measurement campaign was planned at some river points to obtain in-situ velocities and cross section areas. These measured data were also used as part of another research whose objective was the hydraulic characterisation of the Chaguana River by using the HEC – RAS model (Vivas 2004). The location of measured river points is shown in Figure 3.5. The monitored points were selected mainly on the basis of accessibility. The campaign was conducted in July 2002.

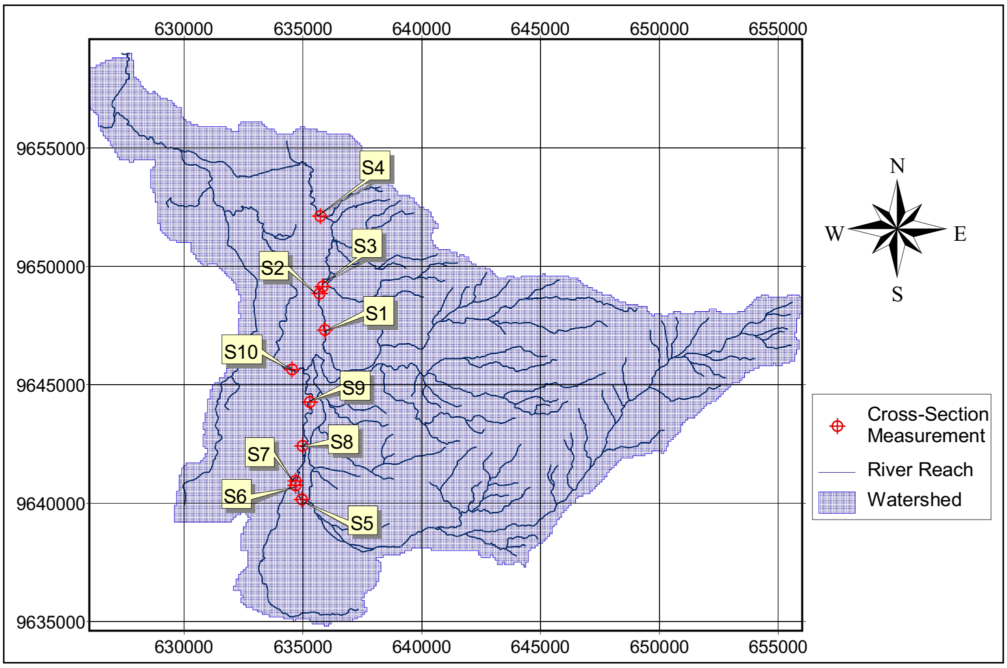


Figure 3.5. Location of measured cross-sections during July 2002

Velocities were measured with a GENERAL OCEANICS Digital Flowmeter, Model No. 2035MKIV. To obtain a representative velocity, the probe was located at one third of the river depth on several places in the same cross section. Then, an average velocity was estimated for the specific cross section. The cross section profiles were obtained by topographic measurements (photo 3.2). Measured data and calculated flows are shown in Table 3.6.

Photo 3.2. Measurement of the cross section S1 in the Zapote river

Table 3.6. Flow estimations at Chaguana river basin during July 2002

River	*Monitored Point*	*Cross Section (m^2)*	*Average Velocity (m/s)*	*Estimated Flow (m^3/s)*
Zapote	S1	1.55	0.320	0.49
	S2	2.03	0.293	0.59
	S3	3.70	0.164	0.61
	S4	3.34	0.226	0.76
Chaguana	S5	2.72	0.222	0.60
	S6	1.23	0.500	0.62
	S7	3.68	0.289	1.07
	S8	1.91	0.651	1.24
	S9	3.03	0.455	1.38
	S10	7.92	0.224	1.77

3.5.3. Watershed Delineation and Other Hydrological Data Needed for Selected Models

Flow data is mainly represented as vector data. However, some hydrological information should be spatially distributed (raster format) for modelling purposes. The majority of those data can be generated from topographical information by using GIS techniques as shown in Figure 3.6.

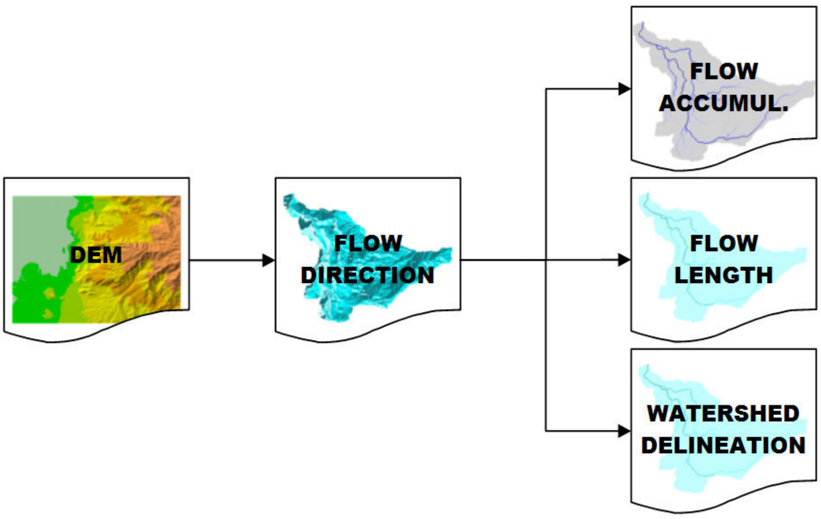


Figure 3.6. Generation of other hydrological data

In this figure, the flow accumulation map generated from the DEM actually shows the cumulative area that drains into a specific cell. By applying colour ramps to the cell values, the map could resemble a river network. However, there are some deviations from the actual river system that can be produced because of:

- **DEM resolution:** The river line in the vector format has actually no real width. The raster cell size is significantly bigger than the visual representation of that line. Thus, the bigger the size of the cell in the DEM, the greater the deviation in the generated river network. In the current research, the cell's size was fixed to 1 hectare as explained in Section 3.4.2.
- **Accurate Topography:** Actual depressions and elevations of the river area may not be well represented by the digital elevation model. Thus, the generated DEM should be adjusted to reflect actual topographical elements at specific sites. There are several methodologies to overcome this problem such as *the burn-in river* method (Maidment 1996). However, such method did not produce good results in the current research. Instead, elevation cell values were changed manually at some areas to force the generated river network to fall within the actual river line.

Figure 3.7 shows a comparison of the same basin spot between the original generated river network and the corrected one by applying the manual adjustments to the Digital Elevation Model. In the figure an extreme case is showed: a conflictive zone where two independent river streams run almost parallel for about 3 km.

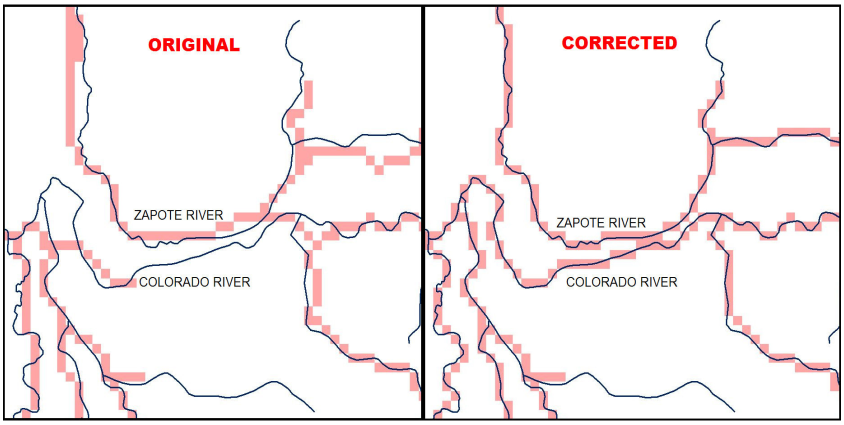


Figure 3.7. Impact of DEM adjustments over river network generation

In the figure, the darker line represents the actual river network while the connected squares represent the generated network. The flow of both rivers (Zapote and Colorado) is from right to left. The original generated network (left) represents the upper part of the Colorado River incorrectly as discharging into the Zapote River. The problem is solved by manual adjustment of the DEM as shown in the corrected river network (right). When modelling non-point sources, this problem must be solved first; otherwise, drainage areas and environmental concentrations could be under- or overestimated.

3.6. METEOROLOGICAL DATA

3.6.1. EXISTING METEOROLOGICAL STATIONS

Meteorological information was gathered among INAMHI[16] weather stations, which are located nearby the Chaguana river basin. However, none of the stations is inside the study site. Historical records on the stations are not continuous, except for one station, and even the recorded period is not the same from one station to another. Table 3.7 shows the nine meteorological stations that potentially could be considered in the river basin and Figure 3.8 depicts the location of the stations related to the watershed. In the table, the distance between each station and the centroid of the basin is also shown. The MACHALA station has the most detailed available information in the area with 27 years of continuous records and measurement of 8 meteorological parameters.

Table 3.7. General description of meteorological stations in the study area

Station	*Distance from Basin's center (Km)*	*Easting (m)*	*Northing (m)*	*Elevation (m)*	*Record Period*	*Type of Data*
Pagua	14.34	636936	9660648	13	1982 – 2000	Monthly precipitation Maximum, minimum and average temperature
Pasaje	16.06	635332	9631874	16	1982 – 1999	Monthly precipitation Maximum, minimum and average temperature
Uzhcurrumi	21.63	657402	9632796	290	1976 – 1999	
Machala - UTM	22.44	621504	9635790	5	1973 – 1999	Monthly precipitation Maximum 24h precipitation Maximum, minimum and average temperature Average relative humidity Sun hours and cloudiness Maximum wind speed and direction
Tenguel	23.25	634815	9669311	10	1965 – 1971 1973 – 1976 1979 – 1980	Monthly precipitation
Tendales	30.89	668850	9633547	750		Monthly precipitation

[16] Instituto Nacional de Meteorología e Hidrología (Ecuadorian National Institute of Meteorology and Hydrology) is the official organization in keeping meteorological records.

Station	Distance from Basin's center (Km)	Easting (m)	Northing (m)	Elevation (m)	Record Period	Type of Data
Huertas	46.49	652199	9601784	1530	1971 – 1977, 1980 1982 – 1986, 1988 1990 – 1991	Monthly precipitation
Santa Isabel	47.51	687718	9638279	1550	1965 – 1987	Monthly precipitation Maximum 24h precipitation Average temperature Average relative humidity Sun hours and cloudiness
La Cuca	49.33	603918	9614373	20	1982 – 1989	Monthly precipitation Minimum and average temperature Sun hours and cloudiness Average relative humidity

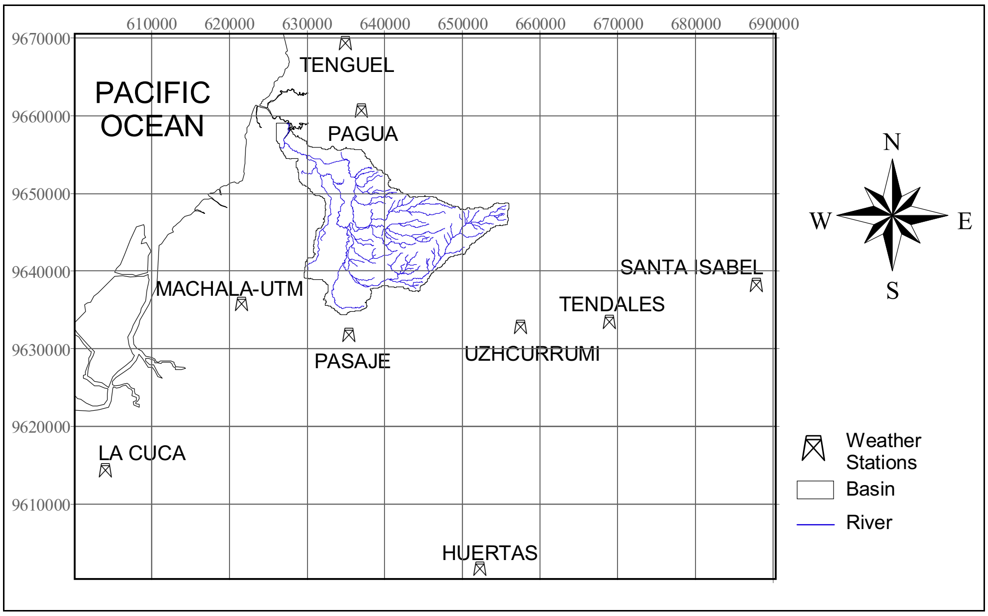


Figure 3.8. Location of available weather stations surrounding the Chaguana basin

3.6.2. MONTHLY PRECIPITATION DATA

The gathered meteorological data were used for the hydrological assessment of the river basin. However, the scope of the analysis is restricted to the average flow condition in the river system within a typical year of precipitation because the only available precipitation data are mean monthly values in all meteorological stations. The existing data is not sufficient at all, though it is the only available data to make the assessment. The actual data show no continuity, and records are not measured during the same periods from one station to another. Thus data selection is a critical step in this basin assessment.

The ideal situation to evaluate a basin's hydrology is to have meteorological stations within the evaluated basin, but that is not always the case. In such cases, some hydrological methods allow the use of data (daily, monthly or yearly data) from nearby stations only, whereas other methods use data from all stations within a certain radius of influence from the basin's centroid. The main drawback of these methods is that the resulting estimated precipitation data is assumed to be the same for the entire basin. In addition, recorded data include both normal and extreme (very high rainfall) events. In order to get a value representative from the whole data set of a station, the median value is used because it is not influenced by very high rain values as the average value is. Finally, the more the stations are geographically distributed, the more the estimated precipitation value is representative for the basin.

For the Chaguana river basin, two stations are nearest to the centroid of the basin (within a radius of 20 km), Pagua and Pasaje (Table 3.7). Their common period of data recording represents 18 years of measurements (1982-1999). Based on these two stations, Table 3.8 shows the mean (median) monthly values estimated for the entire basin. The total annual estimated precipitation is also shown.

Table 3.8. Mean monthly precipitation (mm) estimated from stations located in a radius of 20 km and data recorded between 1997 and 2000.

Station	***Jan***	***Feb***	***Mar***	***Apr***	***May***	***Jun***	***Jul***	***Aug***	***Sep***	***Oct***	***Nov***	***Dec***	***Annual***
Pagua	185.3	230.9	197.5	121.3	29.0	38.9	42.8	30.9	39.7	53.2	34.4	44.7	1048.4
Pasaje	100.9	223.6	211.3	96.6	27.9	30.6	26.6	26.4	26.8	41.4	37.3	53.4	902.8
MEAN	***143.1***	***227.2***	***204.4***	***109.0***	***28.5***	***34.8***	***34.7***	***28.7***	***33.2***	***47.3***	***35.9***	***49.1***	***975.6***

Now considering a radius of influence of 31 km from the centroid of the basin, 4 additional stations could be evaluated to obtain mean monthly precipitation for the entire basin during the same record periods: Machala, Uzhcurrumi, Tenguel and Tendales. By analyzing the record periods, there are three common periods: 1973 – 1976 (Machala, Tendales and Tenguel Stations), 1976 – 1983 (Machala, Tendales and Uzhcurrumi), and 1982 – 1999 (Pagua, Pasaje, Uzchcurrumi and Machala Stations), as shown in tables 3.9, 3.10 and 3.11 respectively.

Table 3.9. Mean monthly precipitation (mm) estimated from stations located in a radius of 31 km and data recorded in 1973-1976.

Station	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*	*Annual*
Tenguel	110.8	352.7	319.6	102.1	75.9	37.9	43.1	24.9	41.1	54.5	43.7	35.3	1241.6
Machala	74.1	222.2	196.0	37.2	26.2	12.7	9.7	16.1	20.9	21.9	14.3	17.8	669.1
Tendales	72.3	156.0	137.9	74.0	43.6	46.2	17.6	12.8	19.9	10.9	15.9	30.4	637.3
MEAN	***85.8***	***243.6***	***217.8***	***71.1***	***48.6***	***32.3***	***23.5***	***17.9***	***27.3***	***29.1***	***24.6***	***27.8***	***849.3***

Table 3.10. Mean monthly precipitation (mm) estimated from stations located in a radius of 31 km and data recorded in 1976-1983.

Station	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*	*Annual*
Machala	92.4	155.9	106.6	78.2	29.8	47.5	33.9	39.3	36.9	79.3	44.6	57.5	801.8
Uzhcurrumi	87.3	98.9	149.2	76.9	33.6	17.1	14.0	10.6	14.6	17.8	21.2	52.0	593.0
Tendales	91.0	74.4	154.1	69.0	58.0	39.9	27.2	10.0	11.4	30.9	65.3	64.3	695.2
MEAN	***90.2***	***109.7***	***136.6***	***74.7***	***40.5***	***34.8***	***25.0***	***20.0***	***20.9***	***42.7***	***43.7***	***57.9***	***696.6***

Table 3.11. Mean monthly precipitation (mm) estimated from stations located in a radius of 31 km and data recorded in 1982-1999.

Station	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*	*Annual*
Pagua	185.3	230.9	197.5	121.3	29.0	38.9	42.8	30.9	39.7	53.2	34.4	44.7	1048.4
Pasaje	100.9	223.6	211.3	96.6	27.9	30.6	26.6	26.4	26.8	41.4	37.3	53.4	902.8
Machala	183.7	253.9	259.1	110.3	37.3	45.9	42.7	38.0	46.9	69.6	53.3	48.0	1188.5
Uzhcurrumi	117.1	117.8	135.4	99.1	34.9	24.6	15.9	11.9	12.6	23.3	26.2	71.5	689.9
MEAN	***146.7***	***206.5***	***200.8***	***106.8***	***32.3***	***35.0***	***32.0***	***26.8***	***31.5***	***46.9***	***37.8***	***54.4***	***957.4***

By considering a radius of 50 km, three additional stations are available: La Cuca, Huertas and Santa Isabel. The two last ones are located around 1500 meters above sea level; the other one is located around 20 meters. Within this influence circumference, three common periods can be analyzed, which represent more than 4 years of records: 1965 – 1971 / 1973 – 1976 / 1980 (Tenguel and Santa Isabel Stations), 1973 – 1976 / 1980 (Machala, Tenguel, Tendales, Huertas and Santa Isabel Stations), and 1982 – 1986 (Pagua, Pasaje, Machala, Uzhcurrumi, Huertas, La Cuca and Santa Isabel Stations). Tables 3.12, 3.13 and 3.14 show the mean monthly precipitation for each evaluated scenario.

Table 3.12.Mean monthly precipitation (mm) estimated from stations located in a radius of 50 km and data recorded in 1965-1971, 1973-1976 and 1980.

Station	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*	*Annual*
Tenguel	99.0	109.8	103.9	72.6	62.1	31.0	18.6	24.5	41.1	47.9	30.9	29.2	*670.4*
Santa Isabel	48.0	66.3	76.4	58.7	25.0	13.1	4.8	6.1	13.9	23.7	22.2	34.1	*392.2*
MEAN	***73.5***	***88.1***	***90.1***	***65.7***	***43.5***	***22.3***	***11.7***	***15.3***	***27.5***	***35.8***	***26.6***	***31.6***	***531.3***

Table 3.13.Mean monthly precipitation (mm) estimated from stations located in a radius of 50 km and data recorded in 1973-1976 and 1980.

Station	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*	*Annual*
Machala	91.4	182.9	171.8	53.5	40.4	18.1	14.6	16.2	27.7	27.3	18.6	19.0	*681.4*
Tenguel	94.1	291.0	211.8	107.2	60.9	37.9	35.5	23.6	42.6	54.2	43.7	35.3	*1037.8*
Tendales	74.5	145.4	144.9	68.9	62.0	60.7	24.8	13.0	20.5	17.0	21.3	33.6	*686.6*
Huertas	215.4	381.4	530.1	372.6	246.2	89.6	18.1	17.5	49.8	28.1	34.2	117.8	*2100.7*
Sta. Isabel	46.0	110.3	136.4	66.1	76.0	19.4	5.6	26.7	8.3	28.8	24.4	34.1	*582.2*
MEAN	***104.3***	***222.2***	***239.0***	***133.7***	***97.1***	***45.1***	***19.7***	***19.4***	***29.8***	***31.1***	***28.4***	***48.0***	***1017.7***

Table 3.14. Mean monthly precipitation (mm) estimated from stations located in a radius of 50 km and data recorded in 1982-1986.

Station	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*	*Annual*
Pagua	208.0	182.0	276.6	136.1	15.2	103.5	180.7	28.7	54.0	54.0	39.8	38.9	1317.4
Pasaje	58.1	139.4	225.9	142.4	25.3	40.1	22.3	18.1	25.3	43.0	62.6	129.2	931.5
Machala	95.4	229.0	268.7	74.8	31.3	66.6	41.0	36.5	34.5	71.8	57.4	76.2	1083.1
Uzhcurr.	68.1	63.8	143.4	97.9	28.9	14.1	15.4	12.9	15.3	34.5	22.3	103.1	619.7
Huertas	394.8	569.2	318.5	444.0	180.3	23.5	59.8	0.0	29.8	139.9	129.5	368.5	2657.8
La Cuca	27.1	152.3	79.3	31.1	8.2	5.4	6.0	4.6	7.9	9.5	0.7	17.2	349.2
Sta. Isabel	42.6	65.1	74.3	88.2	45.1	1.2	2.2	2.0	16.5	23.7	29.3	58.9	449.1
MEAN	***125.6***	***215.8***	***176.8***	***147.2***	***58.8***	***22.2***	***24.9***	***11.2***	***20.8***	***55.9***	***47.8***	***124.8***	***1031.8***

The tables show how the mean annual precipitation for the basin could vary depending on the number of stations included in the analysis and the number of common recorded climate data. Table 3.15 shows how the estimated precipitation for the basin could vary from high to low events depending on the stations involved in the analysis and the evaluated recording period.

Therefore, the right scenario for the so called "normal" event could be selected based on a probability estimation[17] to have the same or higher annual precipitation than the one in the respective scenario (chance of 50% to have higher events).

Another important aspect of the analysis is the influence of extreme events such "El Niño" which is a period where very heavy rains occur in Ecuador. Normally this recurrent event begins at the middle of one year and ends at the middle of next year. Thus, data are normally influenced during two consecutives years by an "El Niño" event from time to time. Reported "El Niño" events were recorded in 1965-1966, 1968-1970, 1972-1973, 1976-1977, 1982-1983, 1986-1988, 1991-1992, 1994-1995, and 1997-1998 (Villacis *et al.* 2001). Therefore, this extreme event has a recurrence interval of 3 to 6 years. The table also includes some El Niño events in their data range periods for the evaluated scenarios.

Table 3.15. Evaluation of precipitation obtained in different scenarios (involved stations and recorded periods).

Scenario	***Evaluated Weather Stations***	***Period of records measured in weather stations***	***Annual Precipitation (mm/year)***	***Probability to have an event greater or equal***	***Classification of the event***
1	Pagua, Pasaje, Machala, Uzhcurrumi, Huertas, La Cuca, and Santa Isabel	1982 – 1986	1031.77	12.50%	High (humid)
2	Machala, Tenguel, Tendales, Huertas, and Santa Isabel	1973 – 1976, 1980	1017.70	25.00%	⇑
3	Pagua and Pasaje	1982 - 1999	975.58	37.50%	
4	Pagua, Pasaje, Machala, and Uzhcurrumi	1982 – 1999	957.39	50.00%	⇐ Mid (normal)
5	Machala, Tenguel, and Tendales	1973 – 1976	849.31	62.50%	⇓
6	Machala, Uzhcurrumi, and Tendales	1976 – 1983	696.64	75.00%	
7	Tenguel and Santa Isabel	1965 – 1971, 1973 – 1976, 1980	531.3	87.50%	Low (dry)

[17] $P(i) = \frac{i}{n+1}$ [3.2]

where ***i*** is the position in an ascending or descending arrangement of ***n*** values.

Based on the previous analysis, it can be concluded that scenario 4, which represents 18 years of recorded data, could resemble normal precipitation events because there is a 50% chance that the event could be higher or lower. In addition, the weather stations in this scenario are fairly distributed around the basin (Figure 3.8). For that reason, scenario 4 is the best combination of weather stations and recorded period to represent a "normal" precipitation event for the Chaguana river basin.

Map Interpolation Procedures

The monthly precipitation values estimated previously are considered as unique valid values for the entire basin allowing for spatial variability from evaluating weather stations only in an implicit form (the value changes only when another station is selected). In modelling, it is sometimes better to explicitly show spatial variability in parameters such as climatic records. This goal is achieved by interpolating data from meteorological stations. However, in this Ph.D. study, interpolation is restricted to a few methods because there are not so many available data (less than 9 stations). Based on the selected scenario, table 3.11 shows the monthly average precipitation data used for the interpolation process for obtaining a normal monthly precipitation map.

The selection of a suitable interpolation methodology for the available data is critical in order to obtain a non-distorted spatial distribution of any evaluated parameter (Mitas and Mitasova 1999). A wrong interpolation procedure could lead to false distribution patterns in the simulations (Mitasova *et al.* 1996). Based on previous works done elsewhere (Hutchinson and Bichof 1983; Hutchinson 1996; Hutchinson and Corbett 1993), the *SPLINE* interpolation method appears the most recommended for processing the climatic data. To obtain the precipitation raster maps of every month, the *SPLINE* interpolation option in ArcView[18] was applied to the stations selected to produce a normal precipitation event. Therefore, a raster precipitation map was generated for every month in a "normal" year (Figure 3.9). In addition, the total annual precipitation was generated by adding all monthly maps. Table 3.16 shows the average values for each raster map and its spatial standard deviation. From the table and figure, the wettest month in a typical year is February with 205 mm and the driest month is August with 25 mm of average rain in the whole basin.

[18] A weight of zero is selected to produce a basic Thin Plate Smoothing Spline which is the recommended interpolation for climatic events.

Table 3.16. Total monthly average and annual precipitation for Chaguana river basin

	Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*	*Annual*
P_{mean}	152.9	205.2	178.6	113.4	28.2	33.0	34.1	25.2	30.4	41.9	37.0	50.9	*930.6*
Stand. Dev.	19.6	36.5	40.2	4.7	1.8	5.8	6.1	7.1	8.8	13.0	7.7	4.9	*132.3*

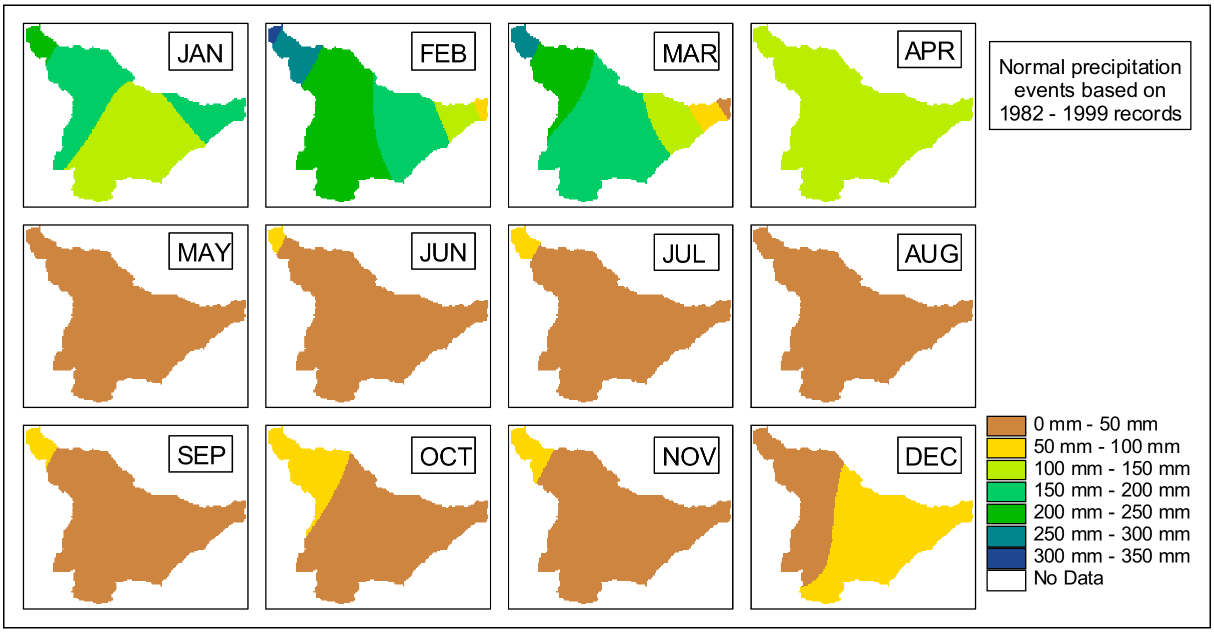


Figure 3.9. Monthly precipitation maps for Chaguana river basin

From figure 3.9, it can be seen that the maps predict more precipitation in the lowlands (west side of the basin) than in the highlands (east side) which might be considered as an uncommon behaviour in a typical precipitation pattern. However, there are two important facts that should be considered when interpreting the interpolated maps:

- The maps do not show the impact of high elevations, because the highest stations (Huertas and Santa Isabel) were not considered in the interpolation procedure. Those stations did not have enough rainfall data to support the interpolation procedure as mentioned before.
- Another important issue is that existing data in the highest stations could be biased as most of their data are influenced by the presence of an El Niño event (extremely high precipitation levels). By reviewing the rain history (interviewing local people) in the areas where the highest stations are, the southern part of Ecuador and the northern part of Peru are very dry (even in the mountain regions).

3.6.3. MAXIMUM 24-HOUR PRECIPITATION

The maximum 24-hour precipitation is a climatic parameter that is useful to obtain other parameters such as the rainfall erosivity factor, peak discharge and so on. In the Chaguana river basin, only three weather stations have records on this parameter: Machala, Pagua and Pasaje. However, these stations do not show the same recording periods for this parameter to produce a representative raster map[19] for a normal event. To overcome this problem, the Gumbel distribution (Benjamin and Cornell 1970) was applied to every monthly value in each station by assuming that these data are recorded continuously through a number N of years.

An EXCEL worksheet was used to evaluate the Gumbel distribution of 24-hour precipitation data in each station. The following equations are used in the calculations.

1. The arithmetic mean 24-hour precipitation (X_{month}) is determined for each month of the year.

$$X_{month} = \frac{\sum_{i=1}^{N} P_{month}(i)}{N} \qquad [3.3]$$

where $P_{month}(i)$ is the 24-hour precipitation for a specific month in year i.

2. The standard deviation (S_{month}) of sample data is determined for each month of the year.

$$S_{month} = \sqrt{\frac{\sum_{i=1}^{N} \left[P_{month}(i) - X_{month}\right]^2}{N-1}} \qquad [3.4]$$

3. Determination of Gumbel distribution fitting parameters: mode (μ) and dispersion (α)

$$\alpha = \frac{\pi}{S_{month}\sqrt{6}} \qquad [3.5]$$

$$\mu = X_{month} - \frac{\gamma}{\alpha} \qquad [3.6]$$

[19] For the interpolation procedure, values used in the stations should represent the same period of measurement.

4. The probability is then established to have the same or greater rain event after a recurrence interval of precipitation (T_{RETURN})[20] occurs. Then, a variate (ω) value is calculated based on this estimated probability.

$$\text{Probability} = \frac{1}{T_{RETURN}} \quad [3.7]$$

$$\omega = -\ln\left[-\ln\left(1 - Probability\right)\right] \quad [3.8]$$

Table 3.17 shows some calculated values of probability and variate for a recurrence interval of precipitation. The return period values shown in the table are the most commonly used in rain analysis.

Table 3.17. Probability – variate values based on a specific recurrence interval of precipitation

Return Period	Probability of event occurrence	Variate (ω)
2 years	0.50	0.3651
5 years	0.20	1.4994
10 years	0.10	2.2504
25 years	0.04	3.1985

5. Finally, a monthly precipitation for a specific recurrence interval is estimated based on the Gumbel fitting parameters for each month (μ,α) and the corresponding variate calculated for that interval. Based on the probability shown in table 3.17, the 2-year precipitation event could be considered as a "normal" event because there is a 50% chance to have greater or lower events.

$$P(i) = \mu + \frac{\omega(i)}{\alpha} \quad [3.9]$$

Tables 3.18, 3.19 and 3.20 show the calculated values of maximum 24-hour precipitation for each month in Machala, Pasaje and Pagua stations respectively. Those predicted values were interpolated to produce raster maps of maximum 24-hour precipitation for each month in every recurrence interval (2 years, 5 years, 10 years and 25 years) as shown in Figure 3.10.

[20] The return period (T_{RETURN}) is the average interval in years between the occurrence of an event of stated magnitude and an equal or more serious event.

Table 3.18. Estimated maximum 24-hour precipitations for each recurrence interval in Machala station

T_{RETURN}	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
2 years	57.1	70.6	60.5	37.0	19.7	14.0	8.7	7.8	8.3	8.5	12.1	25.8
5 years	106.5	114.4	94.3	75.4	41.9	40.1	15.4	16.0	15.8	11.6	29.1	60.3
10 years	139.2	143.5	116.7	100.9	56.6	57.3	19.8	21.5	20.7	13.8	40.3	83.1
25 years	180.5	180.2	145.1	133.0	75.1	79.1	25.4	28.3	27.0	16.4	54.6	111.9

Table 3.19. Estimated maximum 24-hour precipitations for each recurrence interval in Pasaje station

T_{RETURN}	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
2 years	38.1	29.8	51.3	42.7	12.8	9.6	6.1	2.8	5.9	8.1	18.9	45.4
5 years	96.1	55.6	85.1	86.0	35.8	19.7	12.3	3.8	12.2	15.8	39.3	59.6
10 years	134.5	72.7	107.5	114.7	51.0	26.3	16.4	4.5	16.4	20.9	52.8	69.0
25 years	183.0	94.3	135.7	150.9	70.2	34.7	21.5	5.4	21.8	27.3	69.9	80.9

Table 3.20. Estimated maximum 24-hour precipitations for each recurrence interval in Pagua station

T_{RETURN}	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
2 years	56.0	64.9	55.6	46.7	17.7	15.6	14.9	6.2	11.4	16.8	6.4	21.8
5 years	88.9	102.6	100.9	89.2	40.4	39.6	39.7	10.3	37.6	53.8	8.4	55.1
10 years	110.7	127.6	130.9	117.4	55.5	55.4	56.1	13.0	55.0	78.3	9.7	77.2
25 years	138.2	159.2	168.7	152.9	74.6	75.5	76.8	16.5	76.9	109.3	11.4	105.0

This precipitation data were used in the calculation of the rainfall erosivity factor for each station considered in the analysis, but only the 2-year and the 10-year return period which are used in the selected pesticide models.

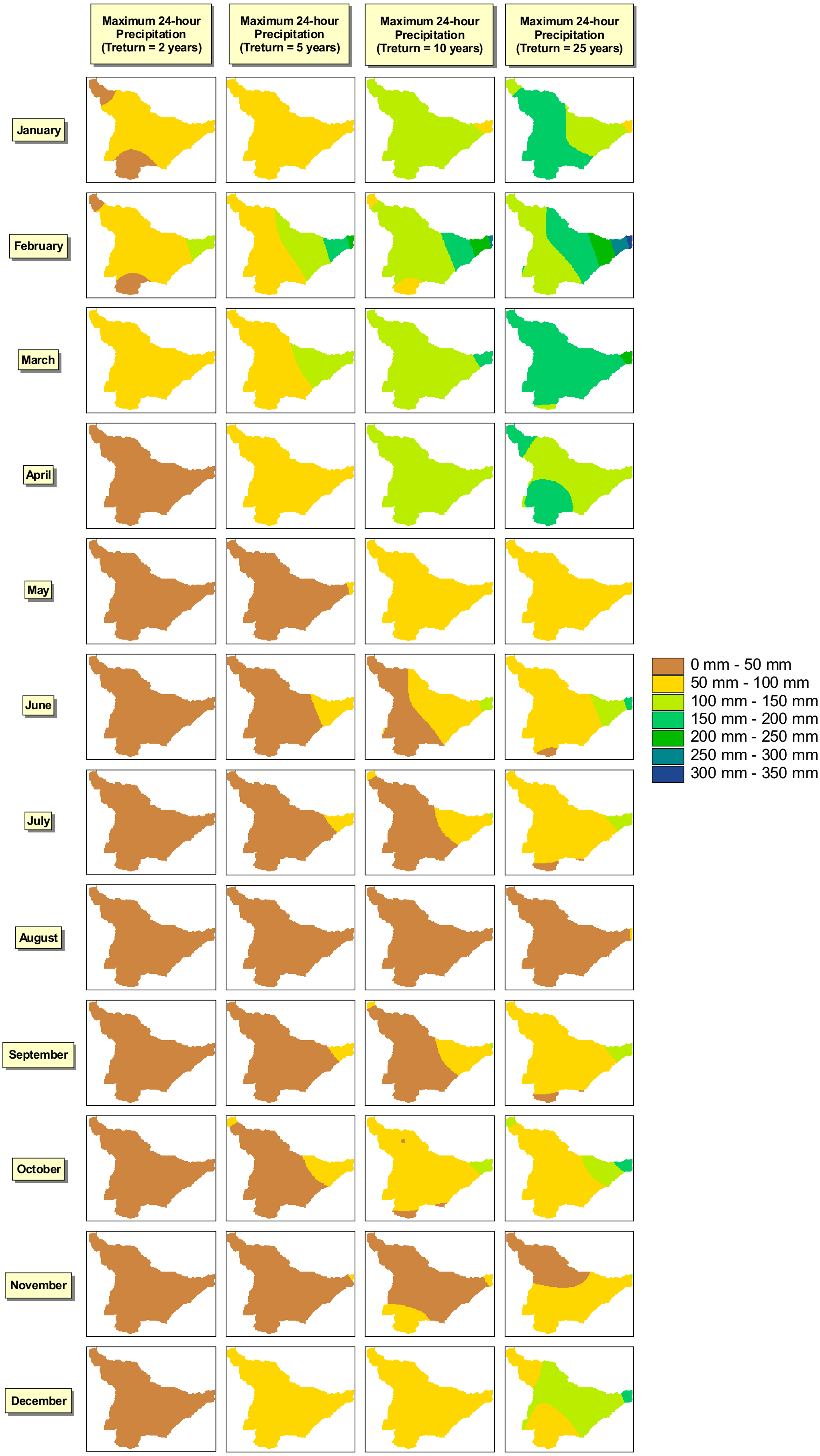


Figure 3.10. Maximum 24-hour precipitation maps for 2-, 5-, 10- and 25-years return periods

3.6.4. AIR AND DEW-POINT TEMPERATURE DATA

Air and dew-point temperature were data required for both pesticide models (SWAT and AGNPS). They can be input as monthly or daily values. Because of the restriction on data availability, only monthly values were generated. Similar to precipitation, air and dew-point temperatures generally show a tendency with topography. In the current assessment, a larger error could be expected because interpolation was done with only three stations (Machala, Pagua and Pasaje). However, this error could be within an acceptable range as explained next. The non-point source pesticide pollution (banana plantations) only affects a part of the catchment area (below 50-m terrain level). Useful stations for temperature evaluation are also located in the same area. Therefore, any error obtained above 50 m level will be irrelevant to the pesticide assessment.

Because the dew-point temperature is not usually monitored in Ecuadorian weather stations, some estimation was performed to obtain that variable. There are several approaches to estimate dew-point temperature such as

- The International Temperature Scale Procedures, established by the International Committee of Weights and Measures (ICWM 1989; Hardy 1998). Although an actual scaling method was proposed in 1990, some equipment could still have scaling methods created in 1968, 1975 or 1976. However, temperature differences between ITSP-90 and ITSP-68 are only in the range of 0.005 and 0.007 °C when the measured temperature is between 20 and 30°C. Therefore, the pesticide assessment will not be affected for the selection of any scale procedure due to this negligible difference in the current assessment context.
- An empirical relationship (Tetens 1930) that uses dry air temperature (T_{AIR}) and relative humidity (RH) to calculate dew-point temperature with ±0.1°C accuracy.
- Some software based on previous formulations such as EZAir (Parks 1998)

In the current research, the following equations based on those methodologies were applied to estimate the dew-point temperature. Those equations are

$$T_{DEW} = -\frac{\left[237.314 \ln\left(\frac{E_T}{6.1078}\right)\right]}{\left[\ln\left(\frac{E_T}{6.1078}\right)\right] - 17.26902} \qquad [3.10]$$

$$E_T = E_W - 0.63\,(T_{AIR} - T_{WET}) \qquad [3.11]$$

$$E_W = 6.112\, e^{\frac{(17.67\,Twet)}{(Twet)+243.5}} \qquad [3.12]$$

Where,

T_{AIR} Maximum Temperature or Dry Air Temperature

T_{WET} Minimum Temperature or Wet Bulb Temperature

T_{DEW} Dew-point Temperature

E_W Saturation Vapour Pressure over water

E_T Actual Vapour Pressure

An ArcView avenue script called Monthly DewPoint Estimation was developed to draw the corresponding raster map based on maximum (dry) and minimum (wet) temperature raster maps. The dry and wet temperature raster maps were produced with a *SPLINE WITH TENSION* interpolation performed with data from Machala, Pagua and Pasaje Stations. Table 3.21 shows the dry air (T_{AIR}) and wet-bulb (T_{WET}) temperature, as mean measurements, used in the interpolation process. The results obtained for the Chaguana basin are also shown in this table. It can be seen that temporal variation through the year on all stations and the basin estimate is quite small (less than 5ºC). In addition the spatial variation in the basin is also small for all months (less than 1.4ºC). Figure 3.11 shows temperature raster maps for a "normal" year.

Table 3.22. Measured maximum and minimum temperatures used to estimate the dew-point temperature map for the Chaguana basin

Station	*Parameter*	*JAN*	*FEB*	*MAR*	*APR*	*MAY*	*JUN*	*JUL*	*AUG*	*SEP*	*OCT*	*NOV*	*DEC*
Machala	T_{AIR} (°C)	30.5	30.6	31.2	31.2	30.2	28.6	26.9	26.6	26.8	26.0	27.2	29.4
	T_{WET} (°C)	21.9	22.3	22.7	22.4	22.1	21.1	19.9	19.9	20.1	20.5	20.5	21.4
Pagua	T_{AIR} (°C)	30.6	30.6	31.4	31.3	30.2	28.5	27.3	26.4	26.9	26.4	27.7	29.4
	T_{WET} (°C)	21.5	21.1	22.0	22.2	21.9	20.8	20.1	19.9	20.1	20.4	20.6	21.2
Pasaje	T_{AIR} (°C)	31.1	31.9	31.9	31.5	30.1	28.1	27.9	27.3	28.5	28.9	30.1	31.6
	T_{WET} (°C)	21.7	22.2	22.4	22.3	21.5	21.1	20.4	20.1	20.3	20.5	20.7	21.1
Estimated values for the basin	T_{AIR} (°C)	30.6	30.6	31.4	31.3	30.2	28.5	27.2	26.5	26.9	26.3	27.6	29.4
	T_{WET} (°C)	21.6	21.4	22.2	22.2	21.9	20.9	20.1	19.9	20.1	20.4	20.6	21.2
	T_{DEW} (°C)	17.6	17.2	18.2	18.4	18.4	17.4	16.7	16.8	16.9	17.7	17.3	17.6

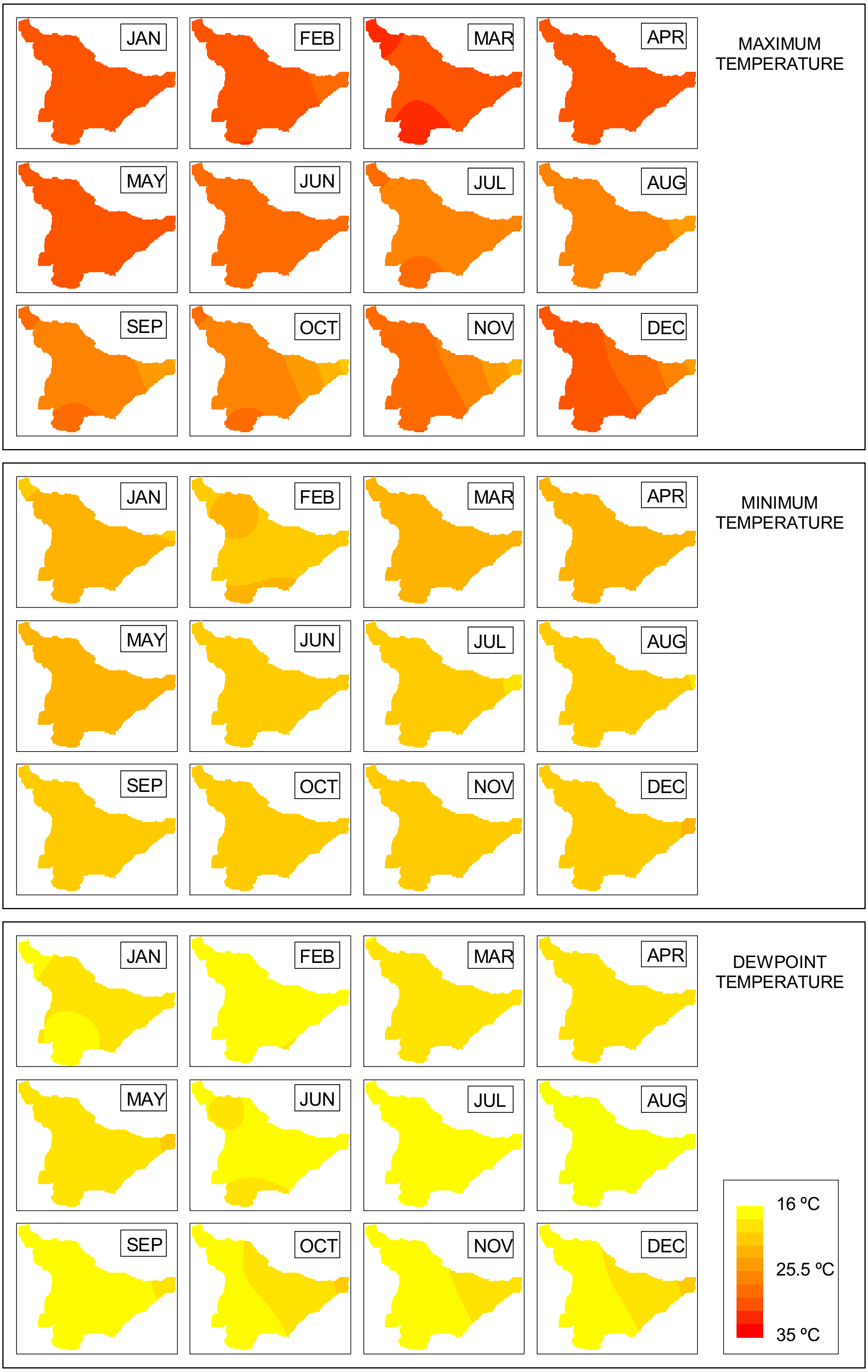


Figure 3.11. Maximum, minimum and dew-point temperature raster maps for a "normal" year

3.6.5. Solar Radiation Data

Solar radiation data are important to account for pesticide degradation due to photolysis. Only SWAT model requires solar radiation as daily or monthly information. In the present study, solar radiation data is lacking in the existing meteorological stations. Only the Machala Station has records on total effective sunshine hours for every month (Table 3.22). From this data, it can be seen that the study area is mostly covered by clouds during the entire year.

Table 3.22. Average total monthly sunshine hours for Machala station

Month	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
Total Sun Hours	65.21	61.54	96.25	93.88	78.43	43.18	40.16	39.62	32.78	21.68	37.64	57.84
Average Hours/day	2.1	2.2	3.1	3.1	2.5	1.4	1.3	1.3	1.1	0.7	1.3	1.9

In order to get solar radiation data for the study area, the following procedure was used

1. ***Gathering information from an existing global on-line database at the NASA Langley Atmospheric Sciences Data Centre:*** This database gives information on solar insolation and other parameters for a 1° grid cell (around 1.2×10^6 Ha) on a 10-year average basis. These data are not necessarily representative for a specific point within the grid cell, and estimated solar data are normally higher than ground measurements (Whitlock *et al.* 2000). In the present study, the Chaguana river basin is located in the grid within 80°E, 3°S, 79°E and 4°S coordinates. Table 3.23 shows the estimated NASA solar radiation on a 10-year average basis for this grid cell.

Table 3.23. Estimated solar radiation from the NASA Langley Atmospheric Sciences Data Center

Month	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
Rad ($MJ/m^2.day$)	16.02	16.85	16.96	16.52	16.09	14.54	15.19	16.60	17.68	17.24	18.14	17.28

2. ***Use of the LARS-WG software tool:*** Due to the overestimation mentioned in the previous paragraph, another procedure is needed to get more representative estimates for the study area. There is one software tool that uses sunshine hour measurements as input to get solar radiation estimates. This software was developed at the University of

Bristol, UK and it is called LARS-WG (Racsko *et al.* 1991; Semenov *et al.* 1998; Semenov and Barrow 1997). This tool is basically a stochastic weather generator and specifically can model solar radiation by using empirical distributions of wet and dry weather series. When there are no solar radiation measurements, sunshine hours are converted to global radiation by means of a regression relationship between these two variables (Rietveld 1978).

However, daily values of sunshine hours are needed to accomplish the weather generation. The input file in LARS-WG needs a daily record of precipitation, minimum and maximum temperature and sunshine hours for the entire period of analysis. In the present case study, only total monthly values for the record period are available. In order to overcome this problem, the following assumptions were made:

- Total monthly sunshine hours were converted to average daily sunshine hours for every month of each monitored year. This value was used as a daily value for every day in the current evaluated month.
- Total monthly precipitation was also converted to average daily precipitation. However, this value was only used as daily value for the total days in a month minus one. Because, the maximum precipitation value in every month was also available, this value was used for that day in the current month.
- In the existing records, temperature values represent the average value in the month, so this value was used as a daily value for every day in the current month.

Table 3.24 shows the average daily solar radiation generated from the Machala station weather data for every month. Figure 3.12 shows how the tendency of the solar radiation curve (left axis) follows the one of the sunshine hours' curve (right axis). Despite the good agreement, it is always better to have some solar radiation measurements to validate these estimations.

Table 3.24. Estimated average monthly solar radiation by using LARS-WG in the Machala station

Month	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
Mean Rad (MJ/m^2.day)	9.7	10.1	12.8	12.3	10.4	7.2	7.2	7.1	7.2	5.8	7.1	8.8

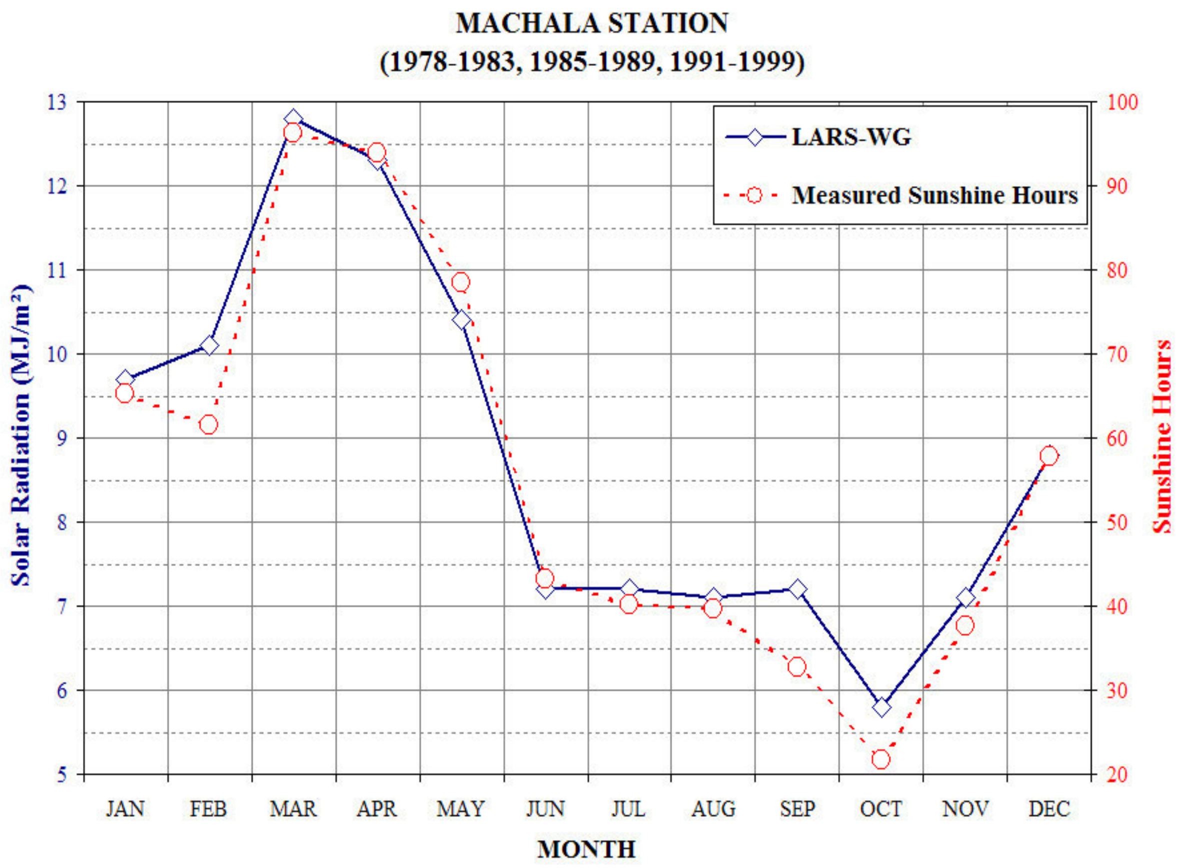


Figure 3.12. Estimated solar radiation and measured sunshine hours in the Machala station

Solar radiation is a parameter that varies not only by the weather conditions but also by the geographic latitude of a specific site. Thus, a solar radiation map would show the spatial variation of this parameter within the map. However, the current study site covers a small geographic spatial variation (less than 0.3° in latitude), which would produce only a slight variation in the parameter values. For that reason, the resulting raster map for the study area will be one with the same value in every raster cell, showing only temporal (monthly) variations and not spatial variations. Values obtained for the Machala station can be accepted as valid solar radiation values for the entire basin.

3.7. SOIL DATA

Another important group of data concerns soil information such as soil texture, permeability, and organic matter. In the Chaguana river basin, the only available information is the edaphology map which shows general information on soil taxonomic units. There is no spatial variability within each taxonomic unit. A sampling campaign was planned to overcome this lack of information.

3.7.1. SOIL SAMPLING CAMPAIGNS

The soil sampling campaign was conducted in March 2001. To plan the campaign, it was necessary to determine the extension of the survey, the parameters to be analysed and the places to be sampled. The following paragraphs will show the methodology applied in the sampling plan.

Sampling Depth and Soil Properties: As previously stated, the available soil data was too general for the selected watershed. Thus, it was necessary to generate more specific soil maps for the study site. In the context of pesticide assessment, the surface soil is more important than deeper soils (Boesten *et al.* 1999; Moorman *et al.* 1999). Thus, the first 50 cm of soil below ground level was evaluated. In addition, soil properties such as soil-water content, bulk density, organic matter and sand-silt-clay content were selected as the main objectives in the soil analysis.

Number of Sampling Sites: The second problem was to answer how many sampling sites are sufficient for the assessment. To solve this problem, the following criteria were considered:

- Both geological and edaphological maps define zones where soils can have very similar properties.
- The locations of banana farms are very important from the pesticide assessment point of view.
- The sampling sites depend on accessibility (existing main roads, consent to enter service roads in the farms). Based on this, the potential surface area to be sampled in the watershed is defined by a variable buffer area not more than 25 meters surrounding the existing and accessible roads.

- Although soil variability is a fact of nature, precision in catchment assessment is not the same as precision in farm assessment. From an agricultural point of view, the largest area representing a soil unit to be sampled is 40 hectares (Jacobsen 1999). From a landscape point of view, around 3 to 6 samples should be taken per landscape unit[21] (Manitoba Government 2000)
- It is necessary to use an adequate statistical technique to determine the minimum number of samples.
- The available budget for the sampling campaign is limited.
- The available time for the whole pesticide assessment is restricted to three years.

Thus, the most appropriate design is the stratified sampling design which has already been used in other types of assessment such as the prediction of organochlorine pesticides in sediments and animal tissues in a river basin (Black *et al.* 2000). For the present study, the sampling size was determined on the basis of statistical analysis (Gilbert 1987). Then, considering the other criteria, the recommended number of stratified sampling sites was fixed to 30 locations.

Considering only the watershed surface area (around 32000 Ha), it apparently seems that very few samples are used to evaluate an important problem in the basin (around one sample per every 1100 Ha). However, the current evaluation seems more appropriate when zoning the sampling points. These can be better seen in Figure 3.13 where sampling points are put over GEOLOGY, LANDUSE, ROAD and EDAPHOLOGY maps. In the road map (Figure 3.13d), it can be seen that a large area is difficult to be sampled due to the lack of roads.

<u>Number of samples per site:</u> Due to soil variability, it was necessary to have more than one soil sample per site. In every sampling site, two soil samples were obtained: a core undisturbed sample and a composite disturbed sample (cross-shaped sampling pattern). A cross-shape sampling pattern represents four sampling points at a certain distance away from the core sample (center of the cross), and separated 90° from each other with respect to the center (Figure 3.14). The radius of the cross-shaped pattern was variable depending on the topography. After collection, the sub-samples were mixed together to form a composite

[21] The landscape unit is a geographical land subdivision based on topography, vegetation, land use, regional context and built infrastructure. According to EPA, it is a designation to identify repeating patterns associated with dominant land uses in an area, and defined by the relative proportions of forest, agriculture, and developed (urban) land cover contained in that area.

sample. The core sample was used to determine the soil bulk density and the composite sample was used to determine the other soil properties.

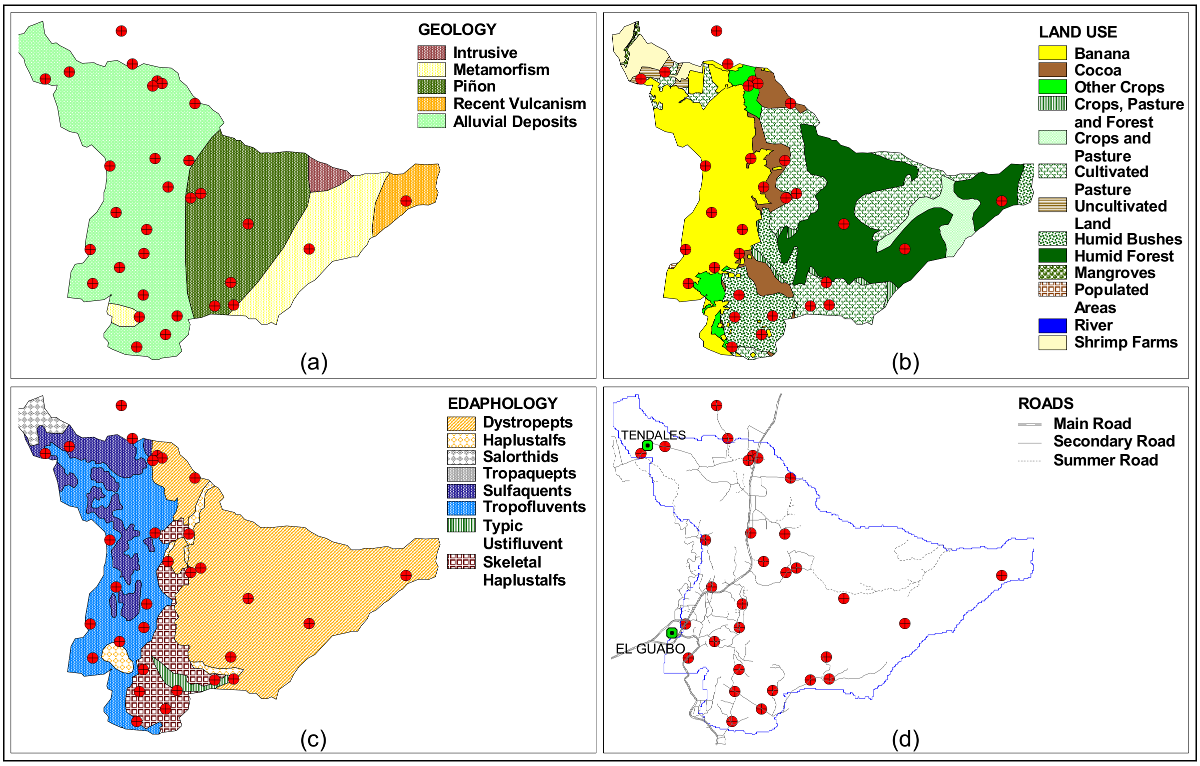


Figure 3.13. Soil sampling point locations compared to (a) Geology Map, (b) Landuse Map, (c) Edaphology Map, and (d) Road Map in the watershed.

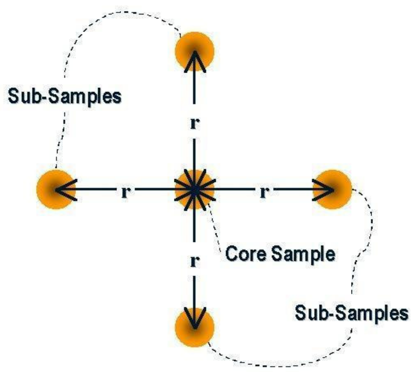


Figure 3.14. Cross-shaped sampling pattern adopted in soil exploration

<u>Issues in sampling site determination:</u> As written before, thirty locations were selected as sampling sites. However, during the sampling campaign, two main factors influenced the decision to obtain only 28 samples:

- Sampling sites were located using available topographical maps made in 1970. However, during the sampling campaign, it was discovered that many roads and access to sampling sites no longer existed or could be accessed only by mules, so it was impossible to get the sampling sites within the available time.
- Many farmers in the basin did not allow entering their properties to take soil samples. They claimed not to be interested in collaborating with the project.

Table 3.25 shows how the final sampling sites were distributed along the different criteria used in the sampling design (GEOLOGY, LANDUSE or EDAPHOLOGY). Zones not important from the project point of view get less sampling weight or no samples at all, mainly based on following considerations:

- Pesticides are mostly used in the banana sector.
- The banana sector is mainly located in alluvial deposits which primarily consist of silty sandy soils.
- From the point of view of edaphology, the banana sector is mainly located in Tropofluvents Soils.

Table 3.25. Zoning of soil samples based on available information for the Chaguana river basin

Information type used in sampling design		*Distribution of samples in a zone*	
Geology	Alluvial Deposits [a]	22	78.57 %
	Cretasic Formation	5	17.86 %
	Metamorphic Formation	1	3.57 %
	Intrusive rock	0	0.00 %
Land use	Banana Sector [a]	15	53.57 %
	Pasture Areas and Other Crops	5	17.86 %
	Hilly Areas	8	28.57 %
	Shrimp Farms Area	0	0.00 %
Edaphology	Tropofluvents [a]	10	35.71 %
	Skeletal Haplustalfs	6	21.43 %
	Dystropets	6	21.43 %
	Tropaquepts	3	10.71 %
	Haplustalfs	2	7.15 %
	Typic Ustifluvents	1	3.57 %
	Salorthids	0	0.00 %

(a) Zones where banana farms are located

3.7.2. LAB ANALYSIS AND RESULTS

Every sample was tested to determine parameters such as organic matter content, soil-water content, bulk density and texture. A summary of test results and some statistics are presented in Table 3.26. In the table, it is also shown the testing methods used during lab analysis. The reader should refer to those methods for further details. Each parameter can be grouped in such a way that a normal distribution can be obtained. For this type of distribution, it is expected that around 68% of the data is found within one standard deviation for the total sampling population.

Table 3.26. Statistical parameters obtained from the analysis results

	%Organic Matter	*% Water Content*	*Bulk Density (kg/m³)*	*% Sand*	*% Silt*	*% Clay*
Testing Method	ASTM-D2974-00	ASTM-D2216-98	ASTM-D2937-00e1	ASTM-D422-63	ASTM-D422-63	ASTM-D422-63
Range	0.13 – 2.01	7.53 – 53.15	1777 – 1078	6 – 99.5	0.4 – 78	0.1 – 73
Mean (X)	1.16	31.11	1409.62	41.3	40.58	18.11
Stand. deviation (s)	0.51	10.44	186.65	24.82	20.92	15.33

When evaluating the stratified sampling, it is desired that every stratum gets a representative number of samples within one standard deviation region and this for each sampling criterion. This statement should hold for every measured parameter. For every parameter data can fall in any of the following regions:

- Region S1: -1 standard deviation (s) to 1 standard deviation (s)
- Region S2: -2s to -1s, and 1s to 2s
- Region S3: -3s to -2s, and 2s to 3s

Table 3.27 shows the sampling distribution within each stratum for every sampling criterion (LANDUSE, GEOLOGY and EDAPHOLOGY). For example, for the LANDUSE Criterion (Banana farms, Hilly areas and Pasture areas; Shrimp farms do not receive any sampling weight), the standard deviation for the organic matter content is 0.51% (Table 3.26). If we consider the Banana Farm Sector in the LANDUSE criterion, it can be seen that around 73% of the 15 samples in that sector fall within the region S1, around 27% fall in region S2, and no samples in region S3. We observe that the distribution of samples fulfils the criteria specified above and that therefore the sampling locations are representative for the type of assessment this study is aimed for, i.e. pesticides in banana plantations.

Table 3.27. Sampling distribution within standard deviation regions for three sampling criteria

SAMPLING CRITERIA			*SOIL PARAMETERS*					
			Organic Matter	*Water Content*	*Bulk Density*	*% Sand*	*% Silt*	*% Clay*
LANDUSE	Banana Farms (15 Samples)	S1	73.3%	73.3%	66.7%	86.7%	86.7%	93.3%
		S2	26.7%	20.0%	33.3%	13.3%	13.3%	0.0%
		S3	0.0%	6.7%	0.0%	0.0%	0.0%	6.7%
	Hilly Areas (8 Samples)	S1	62.5%	25.0%	62.5%	50.0%	50.0%	37.5%
		S2	37.5%	62.5%	25.0%	25.0%	50.0%	62.5%
		S3	0.0%	12.5%	12.5%	25.0%	0.0%	0.0%
	Pasture Areas (5 samples)	S1	60.0%	80.0%	40.0%	60.0%	60.0%	100.0%
		S2	20.0%	20.0%	60.0%	40.0%	40.0%	0.0%
		S3	20.0%	0.0%	0.0%	0.0%	0.0%	0.0%
GEOLOGY	Alluvial Deposits (22 Samples)	S1	68.2%	72.7%	59.1%	77.3%	77.3%	90.9%
		S2	27.3%	22.7%	40.9%	22.7%	22.7%	0.0%
		S3	4.5%	4.5%	0.0%	0.0%	0.0%	9.1%
	Cretasic Formation (5 Samples)	S1	80.0%	20.0%	80.0%	40.0%	40.0%	20.0%
		S2	20.0%	60.0%	0.0%	20.0%	60.0%	80.0%
		S3	0.0%	20.0%	20.0%	40.0%	0.0%	0.0%
	Metamorphic Rock (1 samples)	S1	0.0%	0.0%	0.0%	100.0%	100.0%	100.0%
		S2	100.0%	100.0%	100.0%	0.0%	0.0%	0.0%
		S3	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
EDAPHOLOGY	Tropofluvents (10 Samples)	S1	70.0%	70.0%	60.0%	90.0%	90.0%	100.0%
		S2	30.0%	20.0%	40.0%	10.0%	10.0%	0.0%
		S3	0.0%	10.0%	0.0%	0.0%	0.0%	0.0%
	Skeletal Haplustalfs (6 Samples)	S1	50.0%	50.0%	50.0%	50.0%	66.7%	83.3%
		S2	50.0%	50.0%	33.3%	33.3%	33.3%	16.7%
		S3	0.0%	0.0%	16.7%	16.7%	0.0%	0.0%
	Dystropepts (6 Samples)	S1	100.0%	66.7%	83.3%	50.0%	66.7%	33.3%
		S2	0.0%	16.7%	16.7%	33.3%	33.3%	50.0%
		S3	0.0%	16.7%	0.0%	16.7%	0.0%	16.7%
	Tropaquepts (3 Samples)	S1	66.7%	66.7%	33.3%	100.0%	66.7%	100.0%
		S2	0.0%	33.3%	66.7%	0.0%	33.3%	0.0%
		S3	33.3%	0.0%	0.0%	0.0%	0.0%	0.0%
	Haplustalfs (2 Samples)	S1	0.0%	50.0%	50.0%	50.0%	50.0%	100.0%
		S2	100.0%	50.0%	50.0%	50.0%	50.0%	0.0%
		S3	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Typic Ustifluvents (1 samples)	S1	100.0%	0.0%	100.0%	100.0%	0.0%	0.0%
		S2	0.0%	100.0%	0.0%	0.0%	100.0%	100.0%
		S3	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

3.7.3. SOIL MAPS (DIGITAL DATA) GENERATION

As written before, the purpose of sampling several soil sites was to obtain enough information to develop spatially distributed soil data as raster maps (Figure 3.15). Two methods were used to produce the raster maps depending on the case:

1. Interpolation of soil point data to generate primary soil data maps, and
2. Usage of the Map Calculator macro in ArcView to generate secondary soil information by applying standard equations.

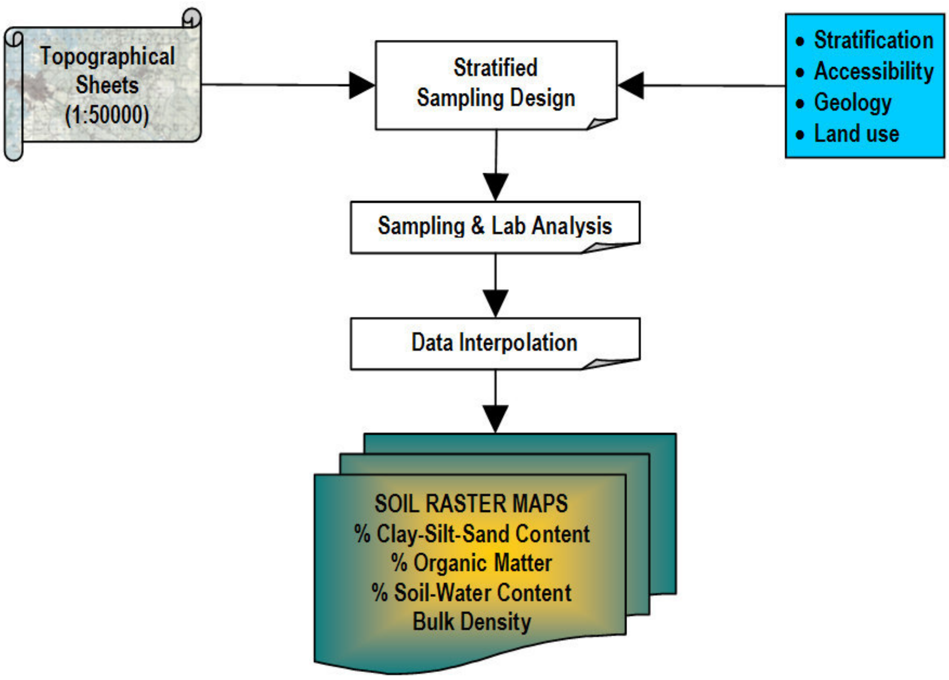


Figure 3.15. Schematic procedure to generate soil maps

Primary soil data: This type of information is related directly to the parameters obtained from lab analysis such as soil texture content (*SAND*, *SILT*, and *CLAY*), organic matter (*ORGMAT*) and water content (*WATCONT*). The soil maps were generated by applying Kriging interpolation, as shown in Figure 3.16. Soil texture maps were corrected in order to have a value of 100% in their arithmetic addition[22]. Each of three texture maps was adjusted by adding or subtracting the mean error value (MEV) to the original texture value.

$$MEV = \frac{100 - (SAND + SILT + CLAY)}{3} \qquad [3.13]$$

[22] Only sand, silt and clay are considered here size fractions of the soil mass.

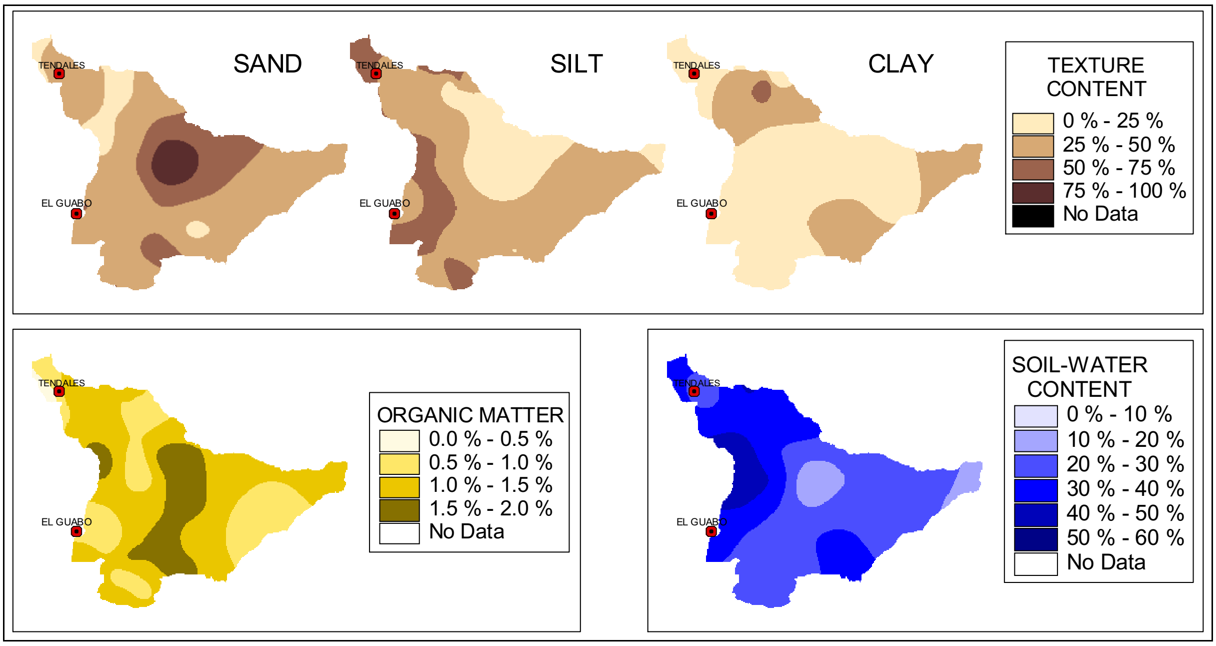


Figure 3.16. Soil primary data maps as a result of sampling interpolation

Regarding the soil water content, this soil property only reflects values obtained during an event (i.e. the sampling campaign). For that reason the generated map cannot be used for other hydrological events. In the present research, this map was used only to show the moisture pattern of the catchment: clayed and fine soils exhibit more water content than sandy soils. In addition, this information is going to be shown as a research by-product to local farmers in irrigation management issues. By showing this information, farmers can see the benefit of having and generating such data.

Regarding the bulk density map, the interpolation procedure did not consider the tillage practices. However, the majority of the crops in the basin are perennial (banana, cocoa and citrics), so tillage practices are almost not a common farming practice in the basin. The soil bulk density is required as an input data in both evaluated pesticide models.

Secondary soil data: Data such as soil albedo, permeability, wilting point, field capacity, and soil classification can be generated with equations or Boolean algebra in the ArcView Map Calculator, as seen in table 3.28. Soil maps are shown in figures 3.17 and 3.18.

In the current research, the soil albedo was determined by using an equation that only considers soil organic matter content. This equation is not completely accurate because the albedo is also function of vegetation cover. The best way to determine soil albedo is by using remote sensing techniques, but these could not be applied in the current research because:

- The basin area is covered by clouds during almost the whole year. This climatic condition makes it impossible to have good satellite images to perform albedo estimations.
- The remote sensing techniques are very expensive and the research did not have enough budget to cover these costs.

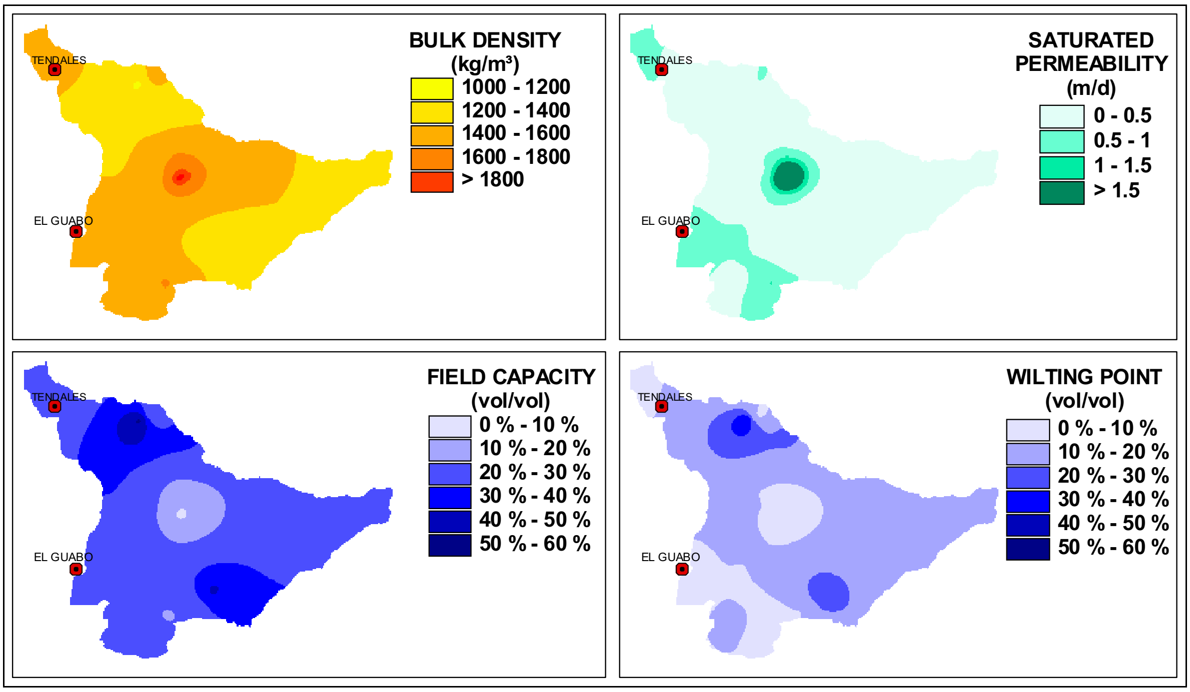


Figure 3.17. Soil properties generated with Saxton's equations

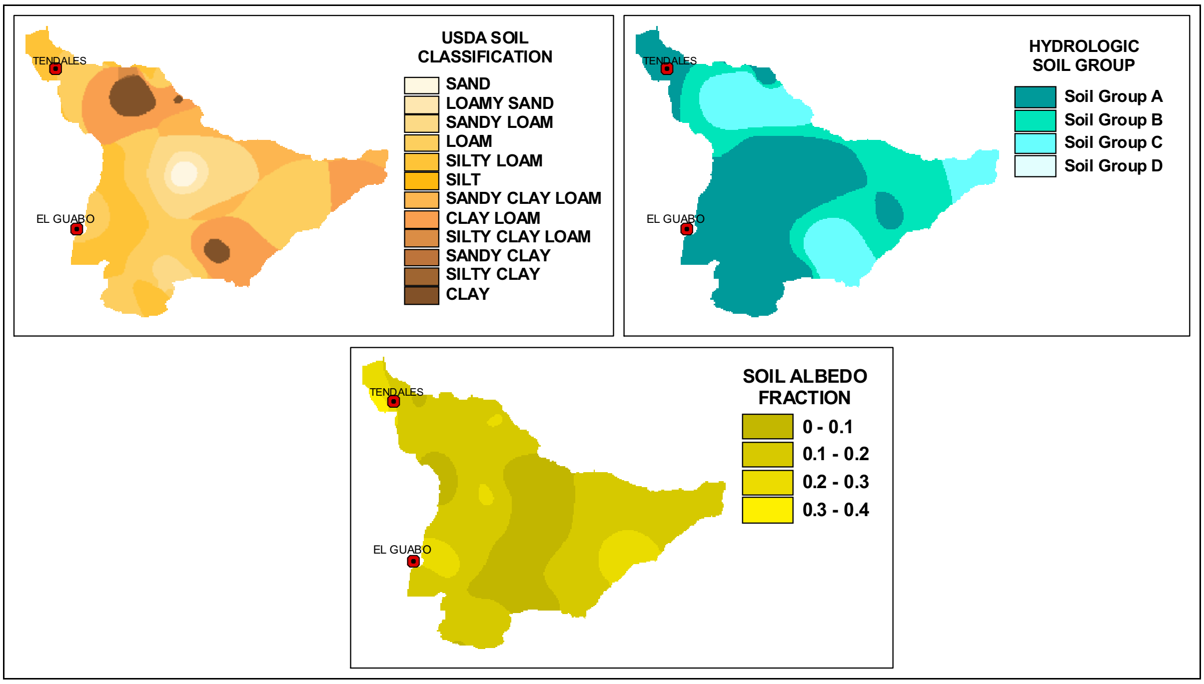


Figure 3.18. Other soil parameters generated with Map Calculator in ArcView

Table 3.28. Secondary data generated with equations and boolean algebra

Data	*Name of the map*	*Equation used*
Soil Properties **(**Saxton *et al.* 1986**):** • **Permeability** • **Bulk Density** • **Wilting Point** • **Field Capacity**	*HYDCOND*	$\exp\left[12.012 - 0.0755\ Sand - \frac{3.895 - 0.03671\,Sand + 0.1103\,Clay - \left(8.7546\times10^{-4}\,Clay^2\right)}{0.332 - \left(7.25\times10^{-4}\ Sand\right) + 0.1276\log(Clay)}\right]$
	BULKDENS	$2.65\left(0.668 + 7.25\times10^{-4}\ SAND - 0.1276\log(CLAY)\right)$
	WILTPOINT	$\left[15.000\exp(4.396 + 0.0715\ CLAY + 4.88\times10^{-4}\ SAND^2 + 4.285\times10^{-5}\ SAND^2\ CLAY)\right]^B$ *where* $B = 3.14 + 0.00222\ CLAY^2 + 0.00003484 SAND^2\ CLAY$
	FIELDCAPAC	$\left[0.3333\exp(4.396 + 0.0715\ CLAY + 4.88\times10^{-4}\ SAND^2 + 4.285\times10^{-5}\ SAND^2\ CLAY)\right]^B$ *where* $B = 3.14 + 0.00222\ CLAY^2 + 0.00003484 SAND^2\ CLAY$
Soil Classification **(**Benham *et al.* 2001**)**	*USDAclass*	*Sand IF* $[(1.5\ CLAY + SILT) < 15]$ *Loamy Sand IF* $[(1.5\ CLAY + SILT) \geq 15\ and\ (2\ CLAY + SILT) < 30]$ *Sandy Loam IF* $\begin{cases}[(CLAY \geq 7)\ and\ (CLAY < 20)\ and\ (SAND > 52)\ and\ (2\ CLAY + SILT) \geq 30] \\ [(CLAY < 7)\ and\ (SILT < 50)\ and\ (2\ CLAY + SILT) \geq 30]\end{cases}$ *Loam IF* $[(CLAY \geq 7)\ and\ (CLAY < 27)\ and\ (SILT \geq 28)\ and\ (SILT < 50)\ and\ (SAND \leq 52)]$ *Silty Loam IF* $\begin{cases}[(SILT \geq 50)\ and\ (CLAY \geq 12)\ and\ (SAND < 27)] \\ [(SILT \geq 50)\ and\ (SILT < 80)\ and\ (CLAY < 12)]\end{cases}$ *Silt IF* $[(SILT \geq 80)\ and\ (CLAY < 12)]$ *Sandy Clay Loam IF* $[(CLAY \geq 20)\ and\ (CLAY < 35)\ and\ (SILT > 28)\ and\ (SAND > 45)]$ *Clay Loam IF* $[(CLAY \geq 27)\ and\ (CLAY < 40)\ and\ (SAND > 20)\ and\ (SAND \leq 45)]$ *Silty Clay Loam IF* $[(CLAY \geq 27)\ and\ (CLAY < 40)\ and\ (SAND \leq 20)]$ *Sandy Clay IF* $[(CLAY \geq 35)\ and\ (SAND > 45)]$ *Silty Clay IF* $[(CLAY \geq 40)\ and\ (SILT \geq 40)]$ *Clay IF* $[(CLAY \geq 40)\ and\ (SAND \leq 45)\ and\ (SILT < 40)]$
Soil Albedo **(**Baumer *et al.* 1994**)**	*ALBEDO*	$\left(0.7\ e^{-0.5596\ ORGMAT}\right)^2$
Hydrologic Soil Groups **(**USDA 1986**)**	*HSG*	Soil Type A if *HYDCOND* is greater than 0.18 m/day Soil Type B if *HYDCOND* is between 0.09 and 0.18 m/day Soil Type C if *HYDCOND* is between 0.03 and 0.09 m/day Soil Type D if *HYDCOND* is less than 0.03 m/day

3.8. LANDUSE DATA

3.8.1. EXISTING LAND USE INFORMATION

Land use data was gathered from the CLIRSEN[23] database. Information in this database is already in ArcView format (shapefiles). There are two types of information in the collected database:

- **Land Cover:** CLIRSEN compiled this data in 1994, and it was digitised from several sources of information such as aerial photographs, satellite information and printed maps at 1:250000 scale. By clipping the basin over the original landuse map, the Chaguana's land cover was obtained (Figure 3.19). Table 3.29 shows the percentage distribution of each activity in the basin. It can be seen that banana is the most important agricultural activity in the basin.

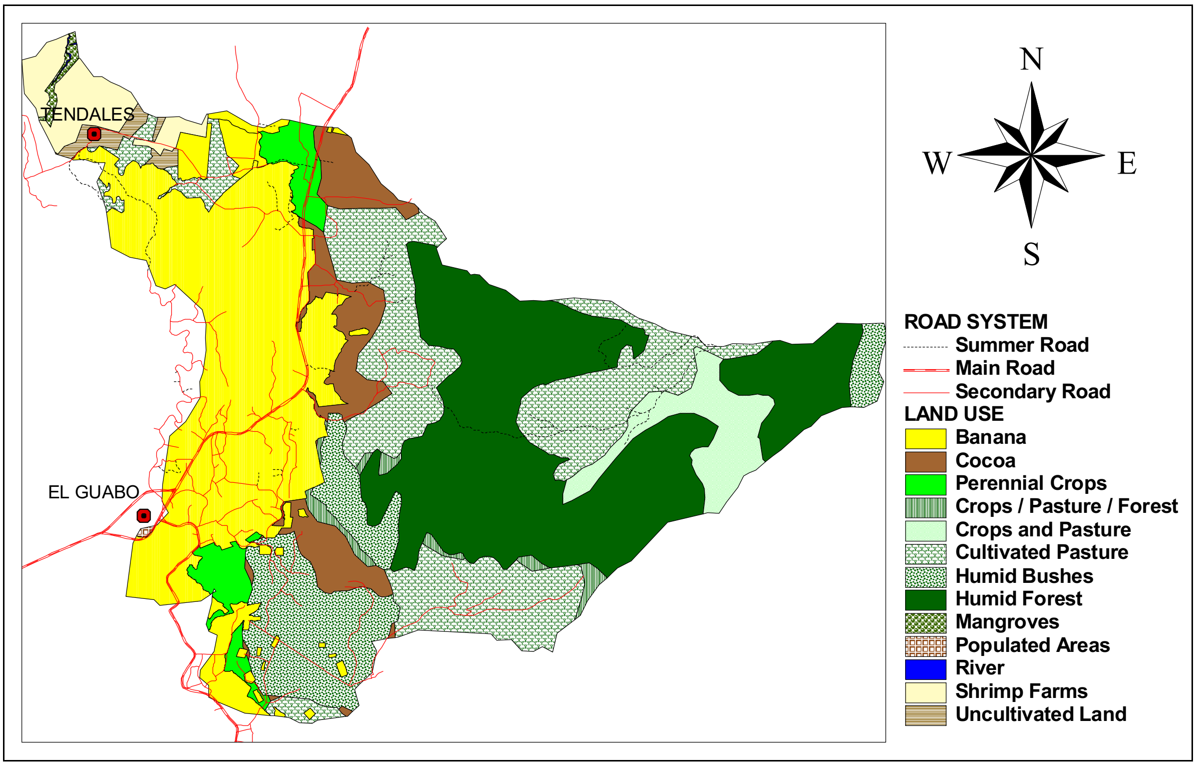


Figure 3.19. Existing land use in the Chaguana river basin

[23] Centro de Levantamiento Integrado y Sensores Remotos (Ecuadorian Centre of Integrated Survey and Remote Sensing) is the official organism to manage land cover and land use data.

Table 3.29. Land use distribution in the Chaguana river basin

Type of Activity	*Land Cover*	*Surface Area*	*Percentage*
Agricultural	Banana crops	8182 Ha	26.21%
	Cultivated Pasture	6177 Ha	19.78%
	Cocoa crops	1927 Ha	6.17%
	Other perennial crops	934 Ha	2.99%
Rangeland	Humid Forest	8142 Ha	26.08%
	Humid Brushes	2950 Ha	9.45%
	Mangroves	86 Ha	0.28%
	Uncultivated lands	261 Ha	0.84%
Mixed Activities	Mixture of crops and pasture	1264 Ha	4.05%
	Mixture of crops, pasture and forest	323 Ha	1.03%
Non-Agricultural	Shrimp Farms	944 Ha	3.02%
	Water	18 Ha	0.06%
	Populated Areas	14 Ha	0.04%

Source: Ecuadorian Centre of Integrated Survey and Remote Sensing (1999)

- **Banana Farms:** There is also information regarding the location and extension of banana farms, and each owner's name. Although this detailed information was not used in the modelling process, it was possible to display a map showing the farms by their extension (Figure 3.20). However, the database is not complete as several banana farms in the southern part of the basin have not been included. The map also shows the river network related to farm locations. It is clear that banana activity mostly affects the lower part of the river basin where other productive activities coexist. This situation can start potential conflicts between users of a river basin as was exposed in the introductory chapter of this thesis.

Based on the map, a farm distribution can be obtained for the Chaguana Basin (Table 3.30). The distribution was arranged in agreement with the classification proposed by the Ecuadorian Ministry of Agriculture for small, medium and big farms (see Chapter 2, table 2.10). By comparing both tables, it can be seen that the farm distribution in the Chaguana basin reflects very well the current situation at national scale conforming this basin as case study. The majority of banana farming comprises small lots with heterogeneous management depending mostly on economics. This situation will be analysed in the next section.

Table 3.30. Banana farm distribution per size in the Chaguana basin

Farm Size	*Percentage*	*Average Size*
Less than 30 Ha	80.2 %	10.6 Ha
30 Ha – 100 Ha	13.8 %	52.2 Ha
More than 100 Ha	6.0 %	186.5 Ha

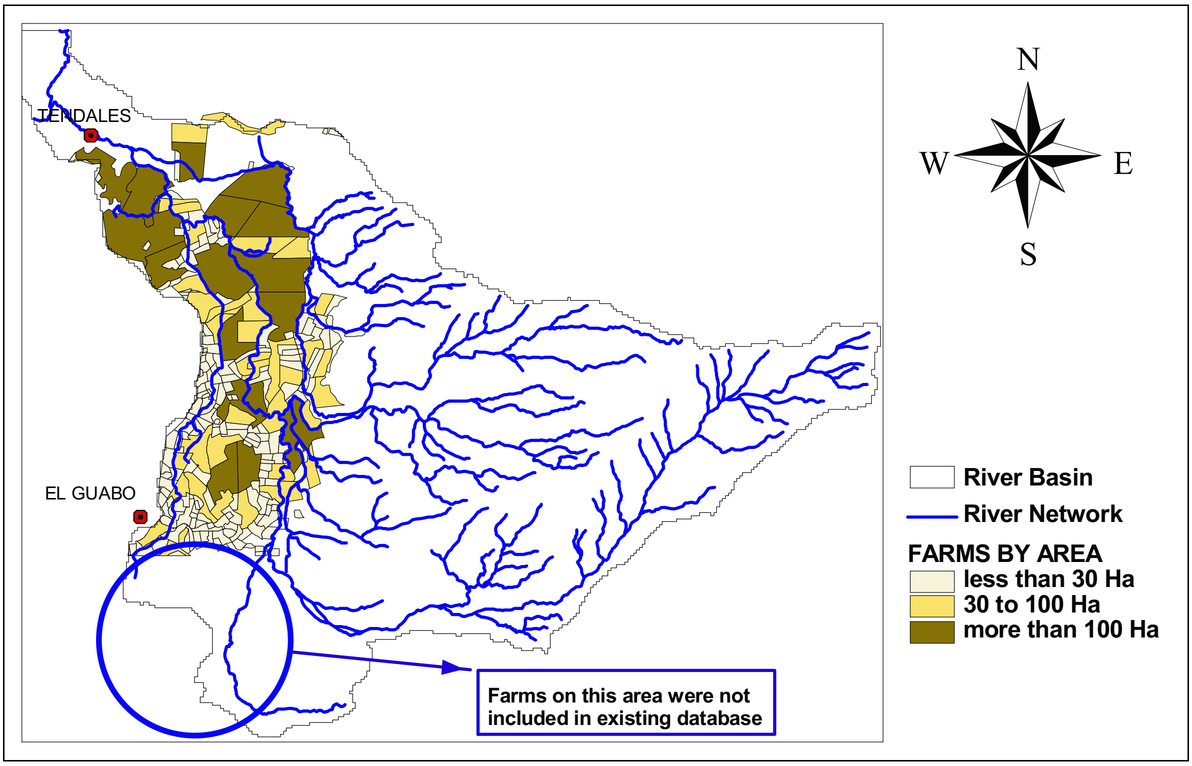


Figure 3.20. Distribution of banana farms by extension in the Chaguana river basin

3.8.2. Pesticide Management Data

As quoted in the previous section, the majority of the banana farms in the basin can be classified as small farms which can strongly affect the way pesticides are being transported inside the basin. Although this problem analysis is out of the scope of the current research, a summary of several findings obtained during the research is presented here.

At the present time, there is a struggle between the banana farmers and the exporters about the in-situ price of a banana box. Farmers claim that prices should be increased because production costs have been increased since the change of the national currency[24]. Among the production costs, the use of agrichemicals represents the highest one after the management costs, as shown in Table 3.31.

[24] In 1999, the Ecuadorian Government decided to use the U.S. dollar as national currency.

Table 3.31. Cost distribution to produce bananas in Ecuador

Cost type	*Percentage*
Management	40.50%
Pesticides and Fertilizers	24.60%
Transport	9.08%
Planting, Caring and Replanting	8.21%
Equipment and others	7.14%
Financial Costs	4.25%
Harvest	3.72%
Irrigation and Drainage	1.57%
Packaging	0.93%

Source: National Association of Banana Producers (1999)

Some small and medium farmers have readjusted their production costs in all items, including pesticide usage, to compensate for the low price of the banana box. On the other side, big farmers can still manage the production costs because sometimes they are also exporters. Therefore, small-medium farmers are forced to use less agrichemicals than the big ones. By using that scenario on the Chaguana basin, small-medium farms are mainly located upstream the river basin while the big farms are downstream. During sampling campaigns, the downstream section of the banana sector showed pesticide concentration values higher than the ones obtained in the upstream section of banana activity (see section 3.11.2).

Thus, the pesticide modelling assessment turns to be complex because pesticides are not used homogeneously in the entire basin, and there is no available information showing the complete pesticide application planning in the watershed. During the research, it was only possible to obtain detailed information regarding the pesticide management in two banana farms: one inside the basin and the other in a different location outside the basin. These data are used together with the general information obtained during the inspections performed on seven banana farms in Ecuador (see Section 2.5). A typical pesticide application planning was structured based on the collected information (Table 3.32).

Table 3.32. Typical application plan in an Ecuadorian banana farm

Application Date	*Julian Day*	*Week*	*Type of pesticide*	*Typical pesticide*	*Application rate*
January 02	2	1	Fungicide	Propiconazole	100 g.a.i. / Ha
January 06	6		Herbicide	Glyphosate	375 g.a.i. / Ha
January 26	26	4	Fungicide	Benomyl	15 g.a.i. / Ha
				Mancozeb	800 g.a.i. / Ha
January 29	29	5	Insecticide	Bacillus Thuringiensis	1 g.a.i. / Ha
February 06	37	6	Herbicide	Glyphosate	375 g.a.i. / Ha
February 15	46	7	Fungicide	Propiconazole	100 g.a.i. / Ha
March 06	65	10	Herbicide	Glyphosate	375 g.a.i. / Ha
March 26	85	13	Insecticide	Bacillus Thuringiensis	1 g.a.i. / Ha
March 28	87		Fungicide	Propiconazole	100 g.a.i. / Ha
April 06	96	14	Herbicide	Glyphosate	375 g.a.i. / Ha
April 08	98		Fungicide	Propiconazole	100 g.a.i. / Ha
April 18	108	16	Mixture	Mancozeb	800 g.a.i. / Ha
				Propiconazole	100 g.a.i. / Ha
May 03	123	18	Fungicide	Propiconazole	100 g.a.i. / Ha
May 29	149	22	Fungicide	Propiconazole	100 g.a.i. / Ha
June 25	176	26	Mixture	Benomyl	15 g.a.i. / Ha
				Mancozeb	800 g.a.i. / Ha
August 7	219	32	Mixture	Benomyl	15 g.a.i. / Ha
				Mancozeb	800 g.a.i. / Ha
August 24	236	34	Fungicide	Propiconazole	100 g.a.i. / Ha
September 15	258	37	Fungicide	Azoxystrobine	100 g.a.i. / Ha
October 04	277	40	Fungicide	Azoxystrobine	100 g.a.i. / Ha
October 25	298	43	Fungicide	Propiconazole	100 g.a.i. / Ha
November 06	310	45	Herbicide	Glyphosate	375 g.a.i. / Ha
November 18	322	46	Fungicide	Benomyl	18 g.a.i. / Ha
November 27	331	48	Fungicide	Propiconazole	100 g.a.i. / Ha
December 06	340	49	Herbicide	Glyphosate	375 g.a.i. / Ha
December 23	357	51	Fungicide	Bitertanol	150 g.a.i / Ha

Based on the table, it is concluded that pesticides are applied on a crop field around 29 times in a year (every 13 days on average). In addition, propiconazole is the most used pesticide in the year with 10 applications through the year (at least one application each month). In the packaging facilities, two pesticides are used after the fruit washing process: thiabendazole and imazalil. Both pesticides are mixed with water and alumina, and they are sprayed over the fruit at a specific location within the packaging facility. The mixture mist is collected in ditches which are connected to the farm channels and then discharged into the river with an average travel distance of 1 km. The pesticide concentration is between 1 and 10 mg of active ingredient per litre for both pesticides. This discharge is produced at least one day per week per farm.

3.9. EROSION DATA

Because pesticide modelling also involves the movement of a chemical attached to suspended particles resulting from erosion mechanisms, it is necessary to also have data related to erosion processes. The majority of pesticide models use the Revised / Universal Soil Loss Equation (RUSLE / USLE). In this equation the following characteristics must be provided:

- The length-slope factor (LS)
- The soil erodibility factor (K_S)
- The rainfall-runoff erosivity factor (R)
- The support practice factor (P)
- The cover management factor (C_M)

These characteristics are discussed below and applied to the current case study.

3.9.1. LENGTH-SLOPE FACTOR (LS)

Basically, the length-slope factor indicates how slope length and slope steepness influence land erosion. It actually measures the relationship between case study topographic conditions to a standard plot conditions (22.12 m of slope length and 9% of slope steepness). The basic L factor equation was first proposed by Wischmeier and Smith (1978), and the first S factor equation was proposed by McCool *et al.* (1987). Combination of these two factors produces the topographic LS factor used in the RUSLE and USLE equations.

The LS factor is usually either estimated or calculated from actual field measurements of length and steepness; however at regional scale, it can be difficult to estimate them because labour-intensive field measurements are obviously not always feasible (Van Remortel *et al.* 2001). For that reason, several authors have proposed methodologies to estimate the LS factor, such as Van Remortel et al. (2001), Desmet and Govers (1996) and Mitosova *et al.* (1996). Those authors recommend not only considering depositional and ridging areas, but also flow concentrated areas in the slope length estimation. Based on those methods, the Length-Slope factor map was obtained as follows by using the ArcView Map Calculator.

Depositional Areas Determination: It has been determined that depositional areas usually begin when the steepness of the evaluated cell is less or equal to a limiting steepness defined by the “concave” area[25]. In ArcView, the concave area can be obtained by using the avenue statement *Curvature*.

CURVATURE = FILLEDEM100.Curvature (Nil, Nil, Nil, Nil)

If the curvature is negative, the cell is concave related to its neighbours; otherwise, the cell is convex. Therefore, depositional areas are the ones where cells are concave. To create the depositional area map, the curvature map is reclassified by assigning a value of 0 to all cells with a negative curvature and a value of 1 to cells with positive curvature.

Concentrated Flow Areas Determination: Another restriction in the LS factor determination is the location of areas where the flow is concentrated (streams, channels, rivers). These areas should be determined by rasterizing the actual river network. However, as explained in section 3.5.3, there are small deviations between the actual river network and the generated river network by the *FlowAccumulation* statement which is the main input for other hydrological determinations. Then, the concentrated flow areas are determined by reclassifying the Flow Accumulation raster into two classes:

- All cells with a cumulative drainage area less than 12 Ha receives a value of 1 which indicates that no water is concentrated, so the cell is susceptible to runoff erosion[26].
- When the cumulative drainage area is greater or equal to 12 Ha, the water is concentrated in the cell, and the chance of deposition is greater than the chance of runoff erosion. Thus, the assigned value is 0.

Slope Length Determination: Once depositional and concentrated flow maps were obtained, a map (named *DEPO_FLOW*) showing all areas excluded for runoff erosion was produced by multiplying both raster maps. Then, the length of slope is determined by using the avenue statement *FlowLength* considering the flow direction raster map (*FLOWDIR*) and the constrained erosion map.

LENGTH_SLOPE = FLOWDIR.FlowLength (DEPO_FLOW, True)

[25] It is the area surrounding the evaluated cell delimited by its eight neighbouring cells.

[26] Based on a digitized river network from topographical maps, it was estimated that an average drainage area of 12 Ha initiates a river stream.

Length Factor (L): The slope length factor is calculated by applying equations developed by Wischmeier and Smith (1978), Foster et al. (1977) and McCool *et al.* (1989).

$$L = \left(\frac{L_{SLOPE}}{22.12848} \right)^{\frac{\beta}{1+\beta}} \qquad [3.14]$$

$$\beta = \frac{Sin\delta}{0.2688\left(Sin\delta\right)^{0.8} + 1.68} \qquad [3.15]$$

Where

L_{SLOPE} Length of the maximum downhill slope

δ Slope angle

Steepness Factor (S): The slope steepness factor is evaluated with the equation proposed by McCool *et al.* (1987).

$$S = \begin{cases} 10.8\ Sin\delta + 0.03 & \rightarrow \quad Slope < 9\% \\ 16.8\ Sin\delta - 0.50 & \rightarrow \quad Slope \geq 9\% \end{cases} \qquad [3.16]$$

Equations 3.14, 3.15 and 3.16 were used in the Map Calculator to produce the raster maps L_FACTOR and S_FACTOR, which were multiplied to obtain the LS_FACTOR map (Figure 3.21a). Table 3.33 shows the range of results obtained from the previous calculations.

Table 3.33. LS factor values for the Chaguana basin

	Length of Slope (m)	*Slope (%)*	*Length Factor*	*Steepness Factor*	*LS Factor*
Maximum value	1228	172	11.64	14.02	111.99
Minimum value	0	0	0	0.03	0.00
Mean	175	20	1.65	2.76	6.21
Standard Deviation	232	22	1.86	3.00	12.45

3.9.2. SOIL ERODIBILITY FACTOR (K_S)

The soil erodibility factor represents the soil loss due to rainfall impact on a unit plot. It is an integrated average annual value which is mainly related to the soil properties. Although it is advisable to measure this value on the field, reliable field data are not always available. However, there are several methods to estimate this value mainly based on soil properties. Based on 225 soils around the world, Shirazi and Boersma (1984) related this value to the mean geometric particle diameter.

$$K_S = 0.0034 + 0.0405 \exp\left[-\frac{1}{2}\left(\frac{1.659 + \log Dg}{0.7101} \right)^2 \right] \quad [3.17]$$

$$Dg = \exp\left[0.01 \left(Cl \ln \Phi_C + St \ln \Phi_{St} + Sn \ln \Phi_S \right) \right] \quad [3.18]$$

Where

K_S	Soil erodibility factor
Dg	Geometric mean particle diameter for a specific soil
Cl, St, Sn	Clay, Silt and Sand fraction respectively
Φ_C	Arithmetic mean diameter of clay size limits (0.001 mm)
Φ_{St}	Arithmetic mean diameter of silt size limits (0.00425 mm)
Φ_S	Arithmetic mean diameter of sand size limits (1.00325 mm)

By applying those equations in map calculator, the K_FACTOR raster map was obtained (Figure 3.21b). Table 3.34 shows the range of obtained values.

Table 3.34. K factor values for the Chaguana basin

	Clay (%)	*Silt (%)*	*Sand (%)*	*K factor (Ton.Ha.h.Ha^{-1}.MJ^{-1}.mm^{-1})*
Maximum value	66.5	76.2	94.3	0.043900
Minimum value	0.4	4.3	8.7	0.007484
Mean	21.6	39.2	39.2	0.038658
Standard Deviation	8.6	12.8	12.1	0.007900

3.9.3. RAINFALL – RUNOFF EROSIVITY FACTOR (R)

The erosivity factor represents the soil loss produced by the influence of the energy of raindrop impacts and a 30-min rainfall intensity. Although this factor can be calculated for a single storm, it is advisable to account for all significant storms occurring through the year. Equations proposed by Brown and Foster (1987) are used in the calculation of the erosivity factor

$$E = 0.29 \left(1 - 0.72\, e^{-0.05\, i} \right) \quad [3.19]$$

$$R_I = E\, I_{30} \quad [3.20]$$

Where

E	Energy produced by the impact of a typical raindrop
i	Intensity of the rain
I_{30}	Maximum 30-min intensity for a specific storm

R_I Rainfall erosivity factor for a specific storm

In the Chaguana basin, there is no information regarding this erosivity factor or 30-min rainfall information. Therefore, it was necessary to estimate the factor based on the maximum 24-h precipitation for the basin during a "normal" year (or 2-year return period as explained in Section 3.6.2). The estimated 2-year maximum precipitations were processed to obtain the annual rainfall erosivity factors for each station (Pagua, Pasaje and Machala). The TR-20[27] methodology was followed to convert the 24-h precipitation into 30-min precipitation, and to estimate the corresponding erosivity factors for each station. The rainfall erosivity map was generated by *SPLINE* Interpolation (Figure 3.21c). In table 3.35, a summary of the generated data is given.

Table 3.35. R factor values for the Chaguana basin

	Annual R factor ($MJ.mm.Ha^{-1}.h^{-1}.year^{-1}$)
Maximum value	2115.80
Minimum value	1067.50
Mean	1384.30
Standard Deviation	249.60

3.9.4. SUPPORT PRACTICE FACTOR (P)

The support practice factor represents the soil loss of a specific crop management practice related to the loss produced by upslope and downslope tillage. It is a dimensionless value ranging between 0 and 1. It is necessary to have very detailed information regarding the crop practice (contouring, stripping, tillage and so on) to estimate this factor.

In the Chaguana basin, the only information available is that bananas are cultivated in lots surrounded by drainage channels resembling terraces. Normally, when the banana plant is harvested, the plant is cut and the plant suckers are allowed to grow. Because banana plantations are permanent crops, no mechanical process affects the soil in the farm. To estimate the P factor values in the basin, equations proposed by Foster and Ferreira (1981) and Foster *et al.* (1997) can be used. Those equations are based on the terrace slope and the terrace length. In banana farms, the terrace length is normally less than 50 meters, and the slope grade is similar to the surface slope.

[27] Technical Release No.20: Computer Program for Project Formulation Hydrology, USDA (1992)

$$P = 1 - B\,(1 - P_Y) \qquad [3.21]$$

$$P_Y = \begin{cases} 0.1\exp(2.4\,s) \rightarrow s < 0.9\% \\ 1 \qquad\qquad\quad\; \rightarrow s \geq 0.9\% \end{cases} \qquad [3.22]$$

Where

s Terrace slope grade or ground slope

B Benefit factor. It is an indicator of the probability to have deposition within the terrace. It varies from 0.5 to 1, and it depends on the terrace length. In the banana farms, this value is assumed to be 0.5.

For the non-crop areas of the basin, the same equation was applied, by considering that there are no terraces in all mountainous and rangeland areas. However, there is still a chance of deposition, so the ground slope is considered for the estimations. Table 3.36 and Figure 3.21 show the generated results for the P factor in the study area.

Table 3.36. P factor values for the Chaguana basin

	P factor
Maximum value	1.00
Minimum value	0.55
Mean	0.87
Standard Deviation	0.19

3.9.5. COVER MANAGEMENT FACTOR (C_M)

The cover management factor is also a dimensionless value ranging between 0 and 1. It represents the soil loss produced by all management activities within a period of time (usually one year). In the RUSLE estimation, this value varies every 15 days in order to cover all management activities within a farm (crop rotation, management schemes, plant growth, etc.). In the USLE estimation, on the other hand, this value is taken as a single average value representing the entire year. For areas such as rangelands or perennial crops, a single value is more appropriate to represent the soil loss rate because the actual cover changes very slowly with time.

In the Chaguana basin, this value will later be a variable to be changed in the model runs to calibrate suspended solids measurements in the river with the sediment loadings produced by the C_M factor (see Chapter 5). However, the initial values, assumed for the basin, depend on the existing land use / land cover conditions. Based on the AGNPS database, Table 3.37

shows recommended C_M factor values for the land cover conditions in the Chaguana basin. However, those values were mainly developed for USA land and crop cover. Therefore, some trial-and- error estimations should be done for banana plantations. Figure 3.21e shows the C_FACTOR map for the Chaguana basin.

Table 3.37. Recommended initial C_M factor values for the Chaguana basin

Land Cover	*Recommended C_M Factor*	
	Average	*Minimum*
Forest	0.03876	0.001
Pasture and brushes	0.06590	0.003
Crops	0.28575	0.170
Mixture of Crops and Pasture	0.17582	0.087
Mixture of Crops, Pasture and Forest	0.13014	0.058
Non crop and barren land	0.60000	0.230

Source: AGNPS model database

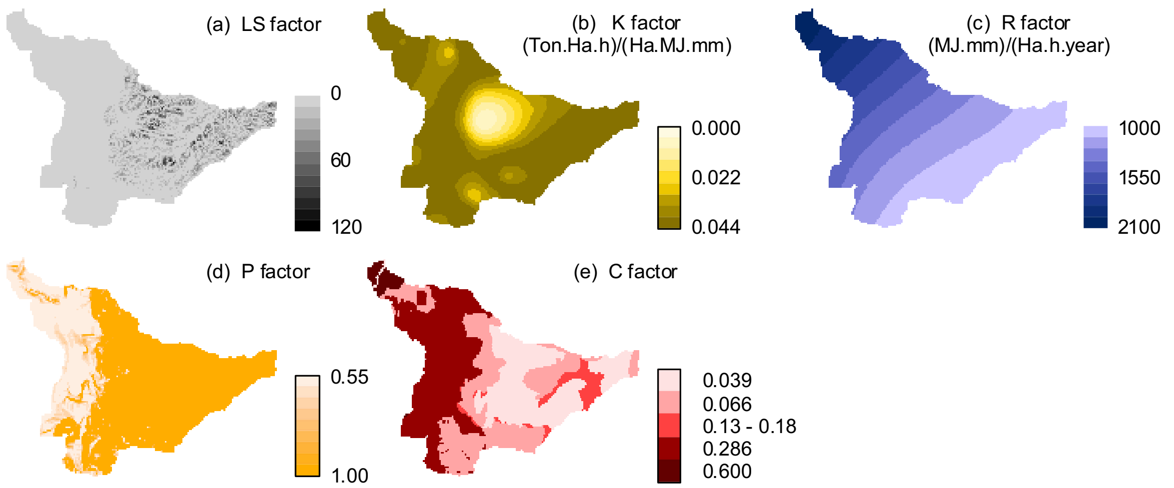


Figure 3.21. RUSLE / USLE factor raster maps for the Chaguana river basin

3.10. CROP DATA

In several pesticide models such as AGNPS and SWAT, it is also necessary to provide information regarding crops to which the pesticides are applied because plants can also affect the way pesticide is transferred to the environment. Time-changing parameters such as canopy cover, crop height and root mass are important for pesticide interception and soil detachment by erosion. Some models such as AGNPS require this information on a 15-day basis. Current models normally have a complete database showing properties of crops mostly cultivated at places where the models were developed. Unfortunately, that is not the case for banana crops.

Data from farmers

In Ecuador, farmers are usually not interested in getting this information on a continuous basis. However, some data which can be used in the evaluated models could be gathered from them:

- The average population density in a banana farm is 1478 plants per hectare which represents around 7 m^2 per plant (diameter of 3 m.)
- The average yield of a banana farm is 27 Tons per Ha.
- The final height of a banana plant differs depending on the plant variety, but the typical height observed during field visits is 3.60 meters.
- The harvest time also depends on cultivated variety ranging from 12 to 15 months. In the current research, it will be assumed that full height and harvesting is reached at the end of 12^{th} month. Once the plant is harvested, it is half-cut so the sucker can feed on it. The farmer can distribute the planting phase in the farm to harvest bananas every week.
- The growing pattern of banana is only available as normalized height values at four growth stages: planting, development, flowering and harvesting (Table 3.38). Normal curves of plant growth show a typical sigmoid shape starting at the origin and tending to a horizontal asymptote. To obtain the height on a 15-day basis, interpolation was done to obtain the curve that fits the data best. In the current research, straight lines were used to fit the data because the height is actually not significantly affecting results in the runoff model. Banana is a perennial crop and farmers do not plant it frequently. Therefore, the plants in banana farms can be considered to be fully developed in the analysis.

Table 3.38. Average growth of banana for the main crop stages

Growth Stage	*Average Julian Day*	*Normalized Height*	*Average Height*
Planting	0	0.01 %	0.05 m
Development	101	75 %	2.70 m
Flowering	245	100 %	3.60 m
Harvesting	365	100 %	3.60 m

Source: Farmer records

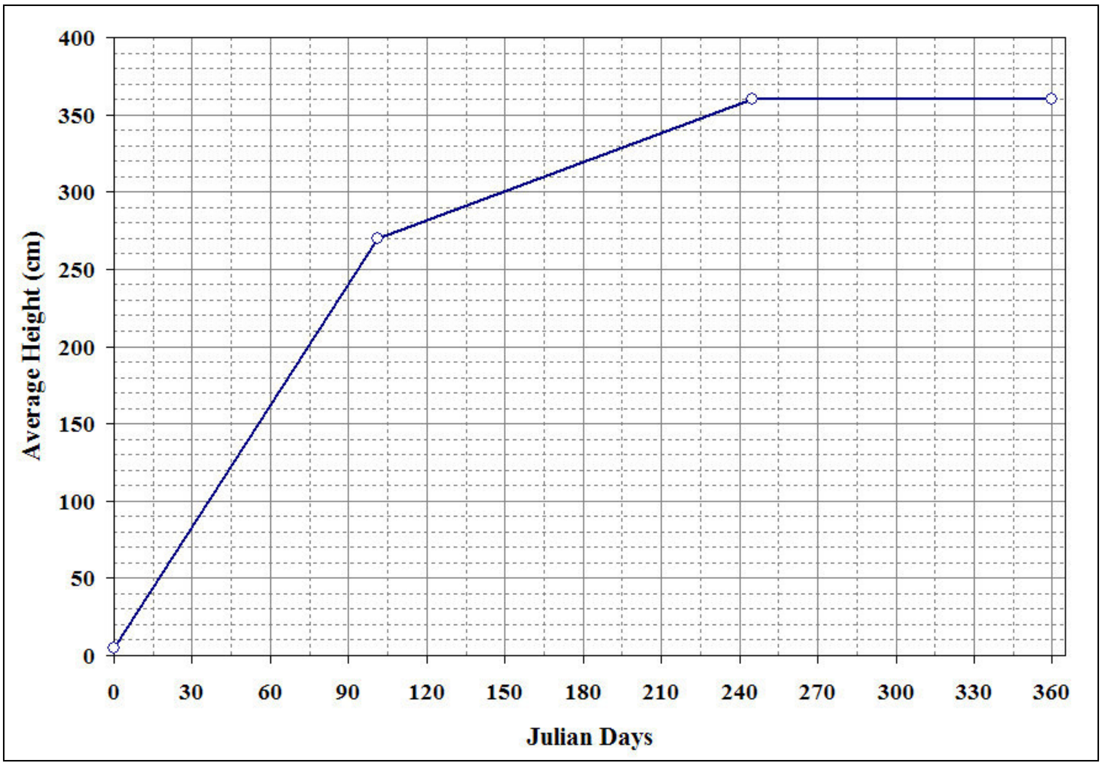


Figure 3.22. Estimated growth of a banana plant

Fresh root mass

The root mass development on a year was estimated based on the research done by Blomme *et al.* (2001a). This study states that root mass can be obtained by using the leaf area, the pseudostem circumference and the tallest sucker length of the banana plant, as shown in equation [3.23]. After searching the literature and some farmer records, values shown in table 3.39 were used to estimate the fresh root mass per hectare of harvested banana plants. For the root development on a 15-day basis, it is assumed that root mass will increase linearly up to the value obtained in the equation.

$$AM_{ROOT}=\left(\frac{0.002066\, A_{LEAF}+0.42659\, \phi_{PSEUDO}+0.171415\, H_{SUCKER}}{1-w}\right)\left(\frac{Den}{1000}\right) \qquad [3.23]$$

where

AM_{ROOT}	Estimated unitary root mass
A_{LEAF}	Mean leaf area of the plant
Φ_{PSEUDO}	Pseudostem circumference
H_{SUCKER}	Height of the tallest sucker
w	Mean water content of the root system
Den	Plant population density in the farm

Table 3.39. Values used for root mass estimation for banana crops

Parameter	*Mean Value*	*Source*
Leaf Area	9900 cm^2	Kumar *et al.* (2002)
Pseudostem Circumference	130 cm	Krauss *et al.* (2001); Blomme *et al.* (2001b)
Height of tallest sucker	85 cm	Blomme *et al.* (2001b)
Root water content	75 %	Several sources

Carbon to Nitrogen Ratio

Another input parameter needed in the AGNPS model is the Carbon to Nitrogen ratio for the banana crop. A study conducted by CEMA[28] (1998), evaluated this ratio in several banana farms at two coastal provinces in Ecuador (Table 3.40). In the current research, the value from the El Oro province is used because that site is the nearest to the study area.

Table 3.40. C:N ratios for banana crops in Ecuador

Province	*Farm Site*	*Mean C:N ratio*
Los Ríos	Quevedo 1	11.75
	Quevedo 2	15.62
	La Mana	13.20
El Oro	Pasaje	17.25

Source: CEMA (1998)

3.11. CHEMICAL DATA

3.11.1. WATER AND SEDIMENT SAMPLING CAMPAIGNS

The Chaguana system has two main rivers: the Zapote and the Chaguana. The Zapote river joins the Chaguana river 5 km before the basin outlet. This basin is not discharging to the sea, but to a bigger basin named the Pagua river basin. In the assessment, the Colorado river, a tributary of the Zapote, is also considered. However, for the purposes of this analysis, it is combined with the Zapote river evaluation as "Colorado & Zapote rivers."

The Chaguana river basin has no historical water quality records, so it was necessary to design and implement a sampling campaign to develop water quality data along the main streams in the catchment. Pesticide concentrations, BOD, TOC, pH, solids content,

[28] Centre of Environmental Studies at ESPOL, Guayaquil, Ecuador

temperature, and dissolved oxygen were measured in the samples, so the existing water conditions were established.

Four measurement campaigns were conducted between November 2001 and November 2002. Each sampling campaign was three days long in order to cover all sampling points. Only one water sample was taken in each sampling point at a certain date and time during the sampling period. Water samples were taken consecutively by following the downstream path of the river once during every sampling campaign. In addition, every certain number of sampling points, a sediment sample was taken for pesticide analysis. The sampling periods were devised in such a way that the climate of a whole year was covered:

1. Transition between dry and rainy season (November 2001)
2. Rainy season (February 2002)
3. Transition between rainy and dry season (June 2002)
4. Dry season (November 2002)

It is important to note that the fourth sampling campaign was influenced by the presence of an extreme climate event (El Niño) which began at the end of 2002 according to INAMHI, the official meteorological organism in Ecuador.

Sampling Methodology

In order to prepare the sampling campaigns optimally, an exploratory field trip was conducted to get direct data from the river system at the moment of the trip. This trip was done in June 2001 which was on the transition between the rainy and dry season. Although most of the observations done during this trip are only representative for the time of the trip (Table 3.41), other non-time dependent information was obtained to improve the future sampling campaigns such as location of banana farms, other land use activities, water discharges and so on. Figure 3.23 shows the travelled path by boat (**thicker line**) during the exploratory trip, and the location of the uppermost place where both rivers are influenced by the tidal effect **(dotted line**). It was impossible to travel by boat further upstream in both rivers because of the low river depth and river barriers[29] (Photo 3.3). Thus, the exploration was continued by following accessible roads adjacent to both rivers (**bold line**).

[29] To pump water into farms, farmers build rock barriers at the downstream side of the pump uptake

Table 3.41. Physical conditions of rivers in June 2001

River	*Average mid-stream depth*	*Average river width*	*Average water velocity*
Chaguana	0.5 m (upstream)	10 m (upstream)	0.2 – 0.6 m/s
	2.0 m (downstream)	30 m (downstream)	
Zapote	0.3 m (upstream)	6 m (upstream)	0.2 – 0.6 m/s
	1.5 m (downstream)	10 m (downstream	

Photo 3.3. Barrier built by a banana farmer to protect a pump uptake

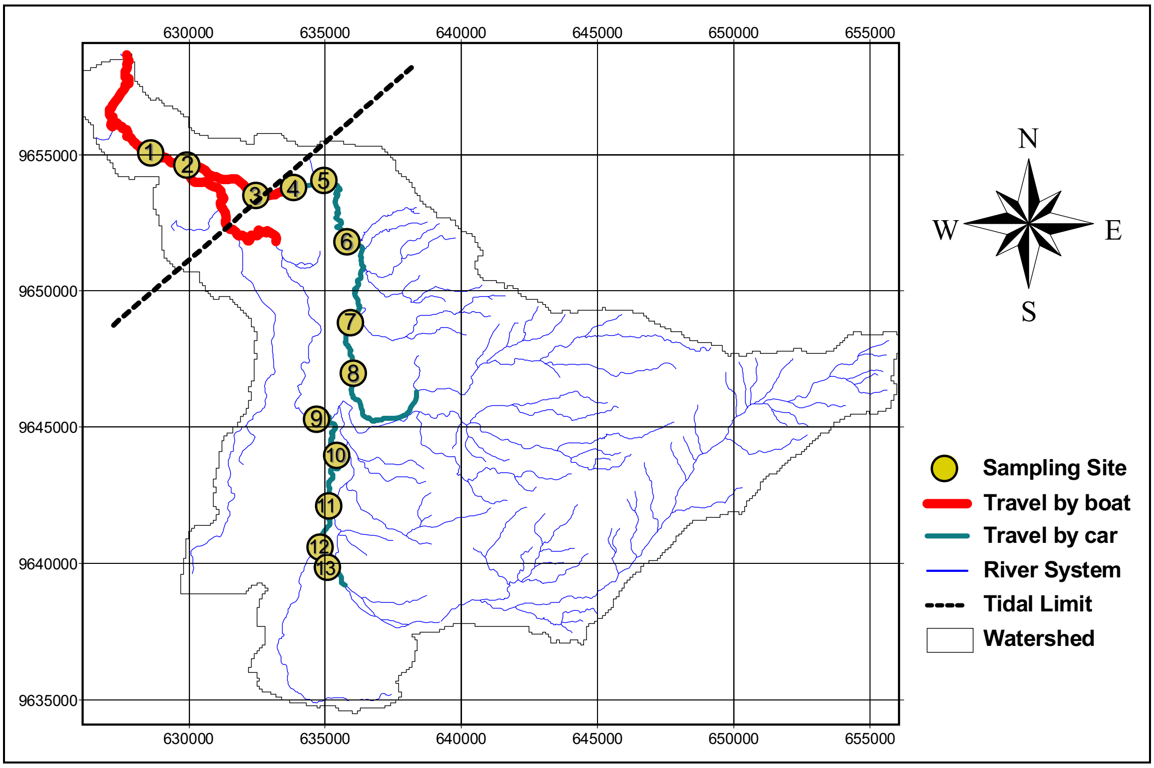


Figure 3.23. Travelling path of the survey done in June 2001 and selected sampling sites

Before the exploratory trip, a sampling procedure called follow-the-plug method[30] was opted for. However, due to the problems mentioned previously, another sampling method was selected: to sample only accessible sampling sites from the most upstream to the most downstream location.

Sampling Sites

Thirteen places which can be accessed at any time of the year were selected to take water samples and eventually sediment samples from each sampling period. From here on, these places will be known as the *Sampling Stations*. However, some other samples were taken at places only accessible by boat during the rainy season or when the river depth allowed to travel by boat. Those places will be called *Reference Sampling Sites* for river stretches not sampled during the whole campaign periods. Table 3.42 shows the location coordinates and a brief description of the sampling stations, which can be seen in Figure 3.23.

Table 3.42. Location of sampling stations in the Chaguana River Basin

Sampling Station	***UTM Coordinates***		***Monitored River***	***Site Description***
	E	***N***		
1	628475	9655363	Main River	Tendales Town near the outlet of the basin
2	629818	9654923		Confluence of both rivers
3	632343	9653798	Zapote	Pump Station at Zapote river
4	633719	9654101		Entrance to “Quirola” farm in the Zapote river
5	634849	9654362		Entrance to “Agricola Leticia” farm in the Zapote river
6	635692	9652118		Main highway bridge over Zapote river
7	635812	9649161		Service road of “Pennsylvania” farm
8	635916	9647300		The most upstream sampling point in the Zapote river
9	634582	9645601	Chaguana	Main highway bridge over the Chaguana river
10	635313	9644289		“La Flores” farm in the Chaguana river
11	635030	9642430		“Juana Fernandez” farm in the Chaguana river
12	634709	9640917		Road bridge to “Chaguana” farm
13	634963	9640170		The most upstream sampling point in the Chaguana river

Measurement Methodology

During sampling, some environmental parameters such as temperature, conductivity and salinity were measured with a YSI-30 instrument. The apparatus probe was introduced up to

[30] It considers sampling the same mass of water while travelling downstream the river.

one third of the river depth at several places across the sampling section. The average value of the measurements is reported for each sampling place.

For pesticide concentrations and other water quality parameters, a sample was obtained in each sampling point. The sample consisted of three 1000-ml bottles (1 glass bottle for pesticide analysis), which were preserved at 4ºC in coolers while they were transported to the lab. Because a sampling campaign was three days long, collected samples were sent to the lab on a daily basis.

Two bottles for each sampling place were analyzed at the Water Quality Lab of the Centre of Aquiculture Services (ESPOL) to obtain Biochemical Oxygen Demand, Total Organic Carbon and Solids Content. Samples in amber glass bottles and sediment samples were analysed in the Ecuadorian Commission of Atomic Energy to obtain pesticide concentration.

3.11.2. Lab Analysis and Results

As said before, there were two groups of data measured during the sampling campaigns: digital data and data obtained from laboratory analysis. In this section, the results obtained will be evaluated and discussed. Results are displayed as colour maps on corresponding river stretches, in which the colour depends to the value assigned to each river stretch. As sampling places are points rather than lines, the following procedure was followed to assign a value to a river stretch:

- A river stretch is connected by two consecutive sampling stations.
- The value assigned to a river stretch corresponds to its upstream sampling station.
- If there are reference sampling sites in the evaluated stretch, these values will be used with the station value to obtain an average for the stretch. In case there is one reference sampling site significantly different from the others, the river stretch will be split at the location of the reference sampling site.

pH

Recorded values of pH during all sampling campaigns were between 6 and 7 which are typical for fresh water. Therefore, there is no significant discharge or no discharge at all of acids or bases that could affect the buffer capacity of the river. The Ecuadorian standard for pH in fresh water rivers is between 6 and 9.

Temperature

The behaviour of water temperature in the sampling points was a little different for all sampling periods. It is important to note that these temperatures are not measured at the same time for each point in both rivers, but they are measured consecutively with increasing time steps from point to point. Therefore, it is better to look at the temperature difference in the whole sampling period for the entire basin rather than interpreting the temperature in each sampling place.

In table 3.43, it can be seen that water temperature varies per sampling campaign. Compared with historical records for those months, the three first sampling campaigns show a normal behaviour. On the other hand, the fourth campaign exhibits a range of temperatures higher than the historical one. This could be explained due to the presence of the initial stages of the 2002 El Niño event. Figure 3.24 shows the measured temperature in the Chaguana river.

Table 3.43. Water temperature difference within the basin per sampling campaign.

Sampling Period	*Temperature Range*	*ΔT spatial*	*Comment*	*Mean $T_{HISTORIC}$*
November 2001	23.5 ºC – 24.6 ºC	1.1ºC	Abnormally dry month	23.4 ºC
March 2002	23.0 ºC – 25.8 ºC	2.8ºC	Rainy month	25.9 ºC
July 2002	22.7 ºC – 24.6 ºC	1.9ºC	Dry month	22.4 ºC
November 2002	24.4 ºC – 27.7 ºC	3.3ºC	Beginning of El Niño event	23.4 ºC

Electrical Conductivity

Conductivity measurements can help in aquatic assessments by providing an estimate of the dissolved ionic matter in the water. This value can be related to the existing environmental state of water bodies: oligotrophic, eutrophic and highly-polluted waters. The two last types of water bodies can exhibit very high values of conductivity. On the other hand, a pollution discharge can be reflected as a sudden change in conductivity. Studies of inland fresh waters indicate that streams supporting good mixed fisheries have a conductivity ranging between 150 and 500 µS/cm (EPA 1997). In the current research, conductivity was measured with digital equipment.

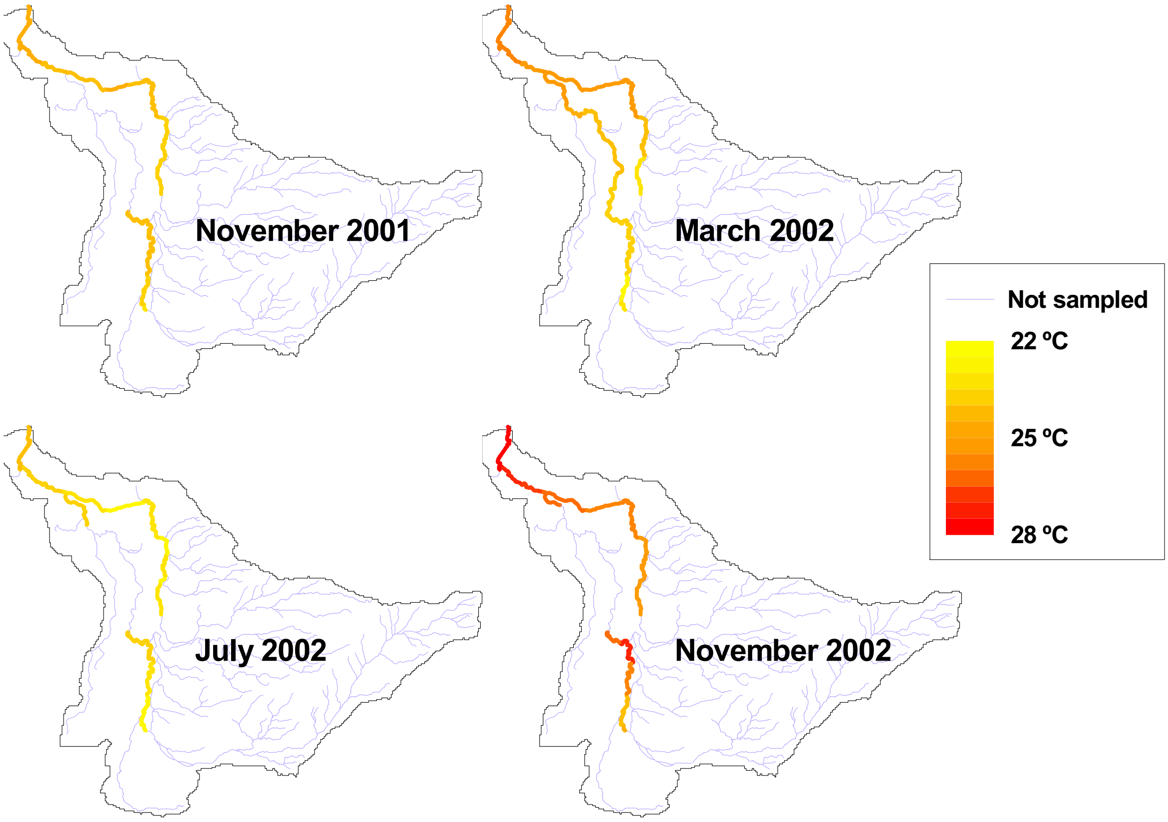


Figure 3.24. Monitored water temperature

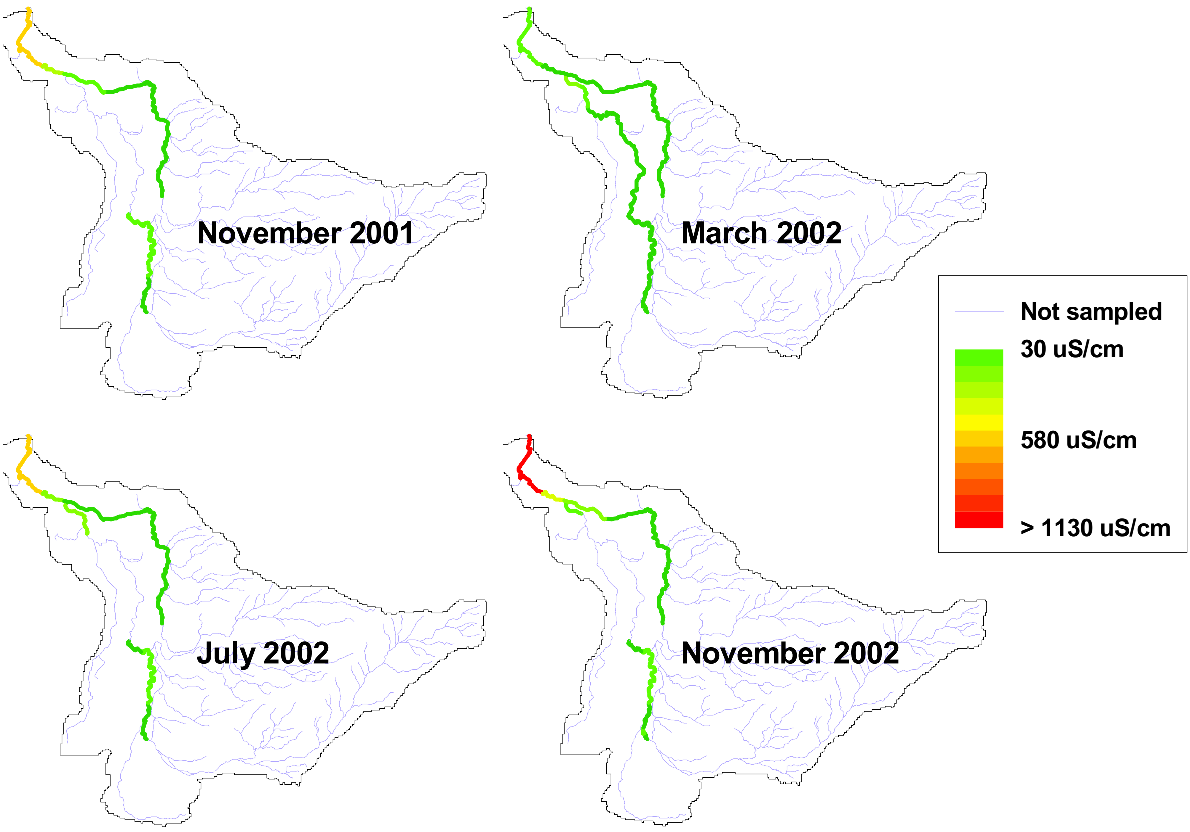


Figure 3.25. Monitored electrical conductivity

In the Chaguana river basin, the majority of the measured conductivity values are below 500 μS/cm, which means that they do not indicate pollution at the sampling dates. However, conductivity values greater than 500 μS/cm were found near the outlet of the basin (Figure 3.25); and this is explained by the potential saline intrusion into the basin. In addition, a town called Tendales (less than 2000 people) is located 5 km upstream from the outlet, and it discharges its wastewater directly to the river.

Solids Content

Suspended and dissolved solids were measured in the laboratory by using STANDARD METHODS AWWA 2540. Results are given in figure 3.26 (Suspended solids) and figure 3.27 (dissolved solids).

The suspended solids content in a river is directly related to the resulting erosion through the basin. Although the Chaguana and the Zapote rivers run through erosive soils (sands and silts), the Chaguana river (left stream) showed more suspended sediments than the Zapote river (right stream) during sampling campaigns, mainly in November 2001 and March 2002. During that period, both river sides of the Chaguana River were under civil works maintenance done by the farms adjacent to the river. Thus, some resulting river erosion was produced at that time. Farmers perform these maintenance works at least once a year.

In river waters, dissolved solids content consist of calcium, chlorides, nitrate, phosphorus, and other ions particles that will pass through a filter with pores of around 2 microns (0.0002 cm) in size. There are many sources where those ions come from; however, pesticides and fertilizers are the main suppliers of dissolved solids in basins where agriculture is the main activity.

Figure 3.27 shows the dissolved solids content measured during each sampling campaign. It is seen that both rivers exhibit very low dissolved solids content (less than 100 mg/l). However, high values (larger than 500 mg/l) were recorded near the basin's outlet. This behaviour correlates very well with the conductivity measurement done at the same places (figure 3.25). As explained before, this last river stretch is influenced by marine intrusion as flood tide enters the basin.

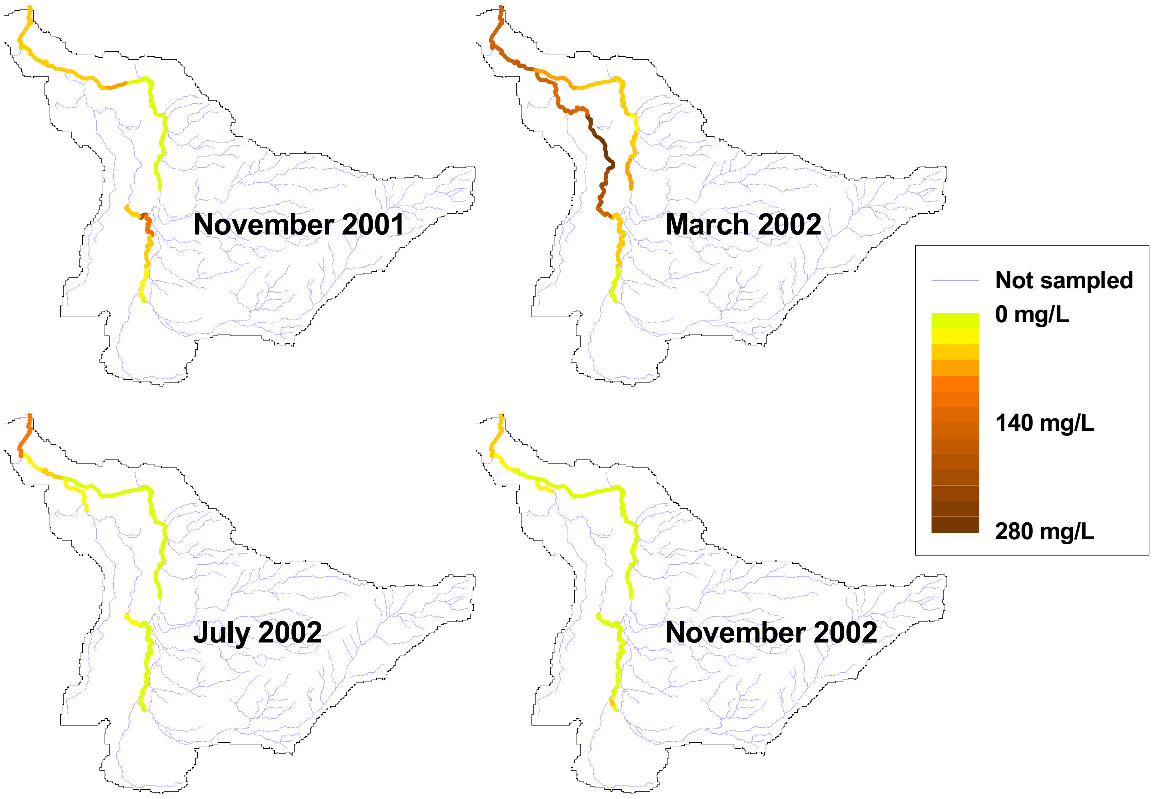


Figure 3.26. Monitored Suspended Solids

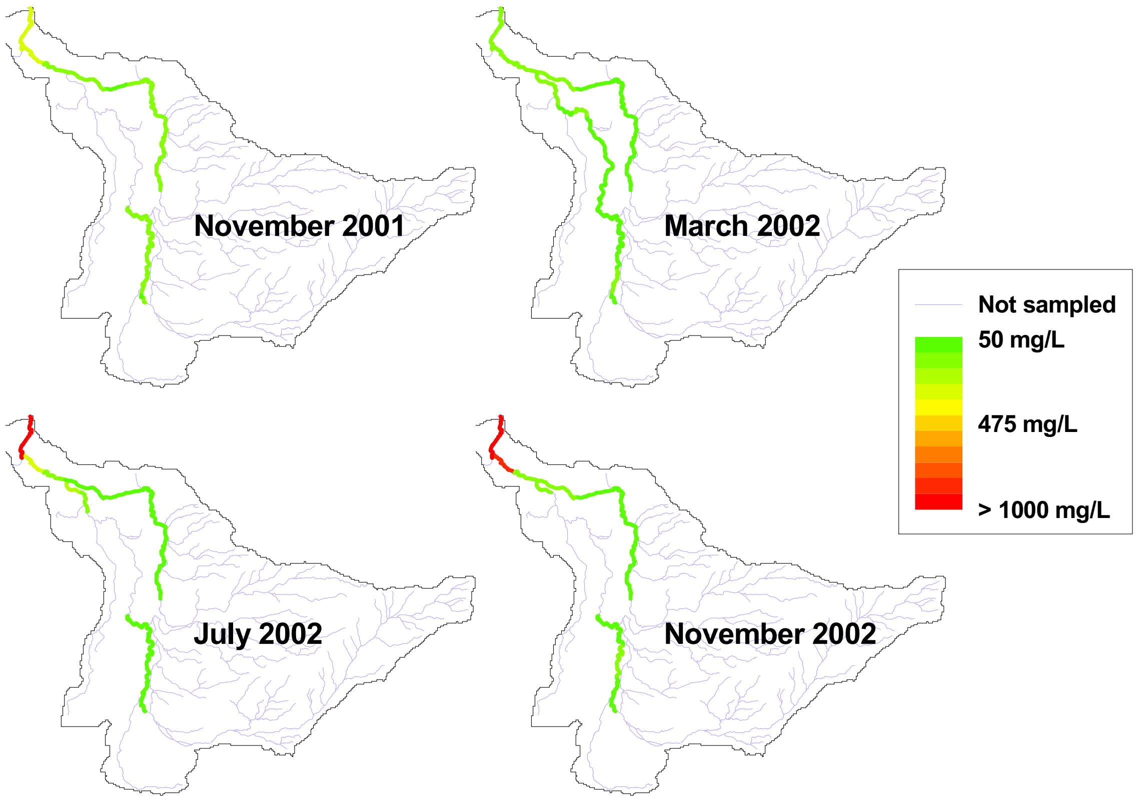


Figure 3.27. Monitored Dissolved Solids

Biochemical Oxygen Demand

BOD measurements are indirect indicators of the amount of organic matter present in water bodies, mainly produced by domestic wastewater discharges. For the case of the Chaguana river basin, the recorded BOD values were between 1 to 4 mg/l which indicates that organic matter discharges into the river were not the main concern for the ecosystem at the moment of sampling. These values are within the acceptable range for fresh-water bodies.

In the evaluated watershed, there are only two main residential areas located at the most western parts of the basin: Tendales and El Guabo (Figure 3.19). In addition, there are some sparse settlements within the farms. Most of the residential areas do not have wastewater treatment systems.

The total amount of people living in the basin is less than 20000, which represents around 23 litres of wastewater per second; and a BOD loading of 500 kg/day or 10 kg BOD per day per kilometre of the monitored stream[31].

[31] The equations used to calculate these values are

$$Q_{sewage} = \frac{N\, q_u}{86400} \qquad [3.24]$$

$$Load_{daily} = \frac{54}{625}\, BOD_{domestic}\, Q_{sewage} \qquad [3.25]$$

$$l_{daily} = \frac{Load_{daily}}{L_{river}} \qquad [3.26]$$

Where

Q_{sewage}	Domestic wastewater flow
q_u	Unitary sewage production which is assumed to be 100 l/p/day in Ecuadorian rural areas.
$BOD_{domestic}$	BOD concentration for domestic wastewater which has an average value of 250 mg/l in Ecuadorian rural areas
$Load_{daily}$	BOD loading on daily basis
l_{daily}	Average linear distribution of BOD loading along the river length
L_{river}	Length of monitored river stream, which in the current research is around 50 km.
N	Number of inhabitants

Pesticides

Water and sediment samples were analysed to obtain pesticide concentrations. As explained in previous sections, propiconazole, imazalil, thiabendazole, glyphosate and tridemorph are the pesticides most used in Ecuadorian banana farms. However, glyphosate and tridemorph could not be traced due to existing lab restrictions in Ecuador (see section 2.5.3); and the others were detected up to 0.01 ppb (minimum detection limit). As expected, a sampling campaign can only detect pesticide concentrations when the spraying is performed within a period of time before the campaign takes place. This is mainly because the river travel time in the basin is less than 24 hours.

Propiconazole is a pesticide applied by air spraying, and it is used mainly for Sigatoka control. During the sampling campaigns some river stretches showed propiconazole concentrations while others did not show any at all (Figure 3.29). From the figure, it is clear that most of the farmers in the basin sprayed their plantations during the March campaign because Sigatoka tends to be present on banana leaves during the rainy season (extreme humid conditions). During the other campaigns, propiconazole is only detected near the basin outlet which is mainly a shrimp area; however, this pesticide could have come from the most downstream banana farm which is located around 7 km upstream the outlet.

Imazalil and thiabendazole are pesticides used during the packaging process of banana mainly as a pesticide mixture. As explained in the *Pesticide Management Data* section 3.8.2, farmers harvest banana at least one day every week. However, the packaging process is performed independently, depending on each farm's operational plan. Detected concentrations of both pesticides can be seen in figures 3.30 and 3.31. The presence of both pesticides in the river is very similar, which is very logical in view of their similar application.

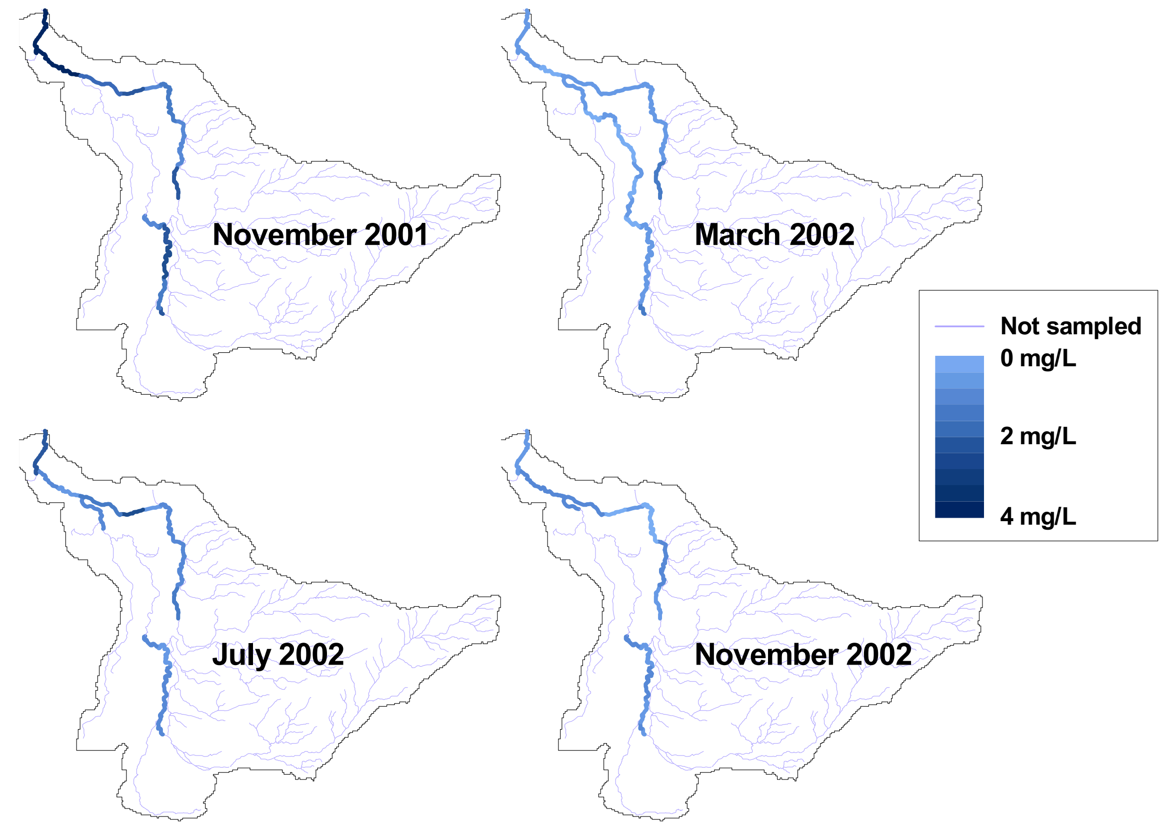


Figure 3.28. Monitored Biochemical Oxygen Demand

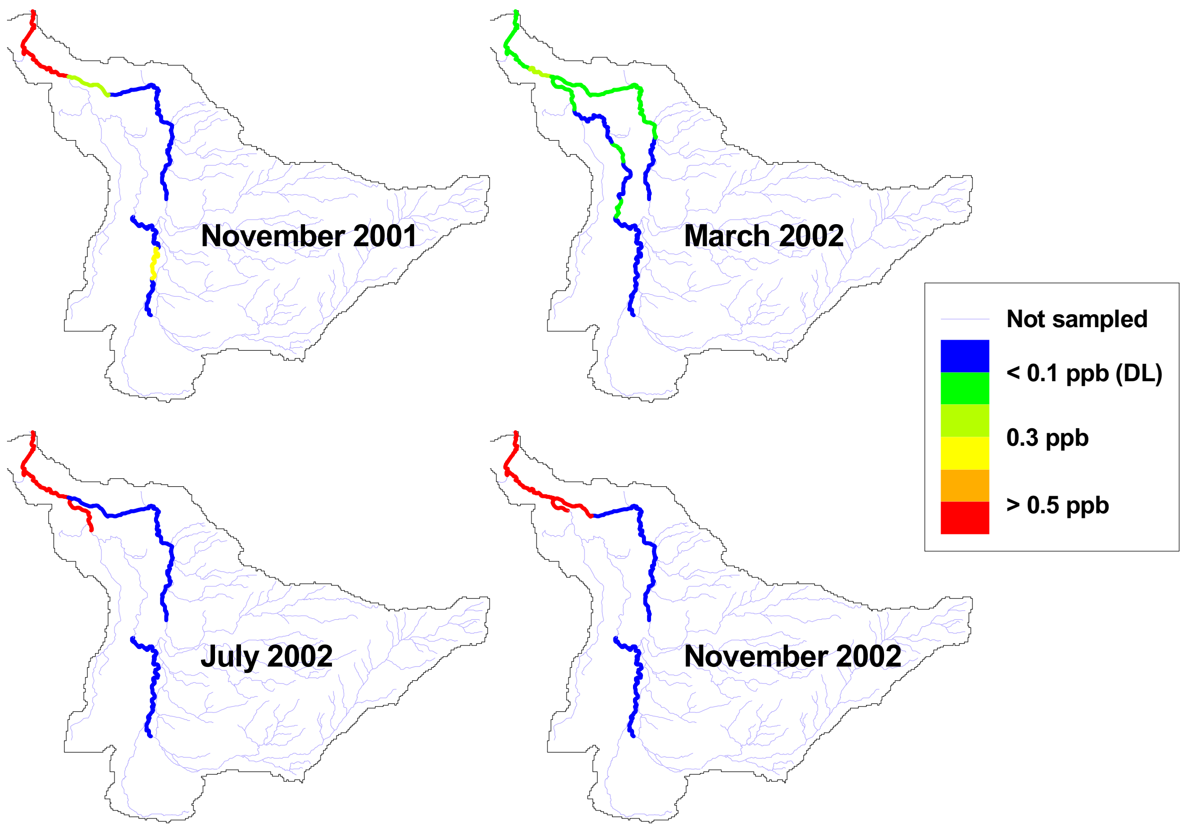


Figure 3.29. Monitored Propiconazole concentrations

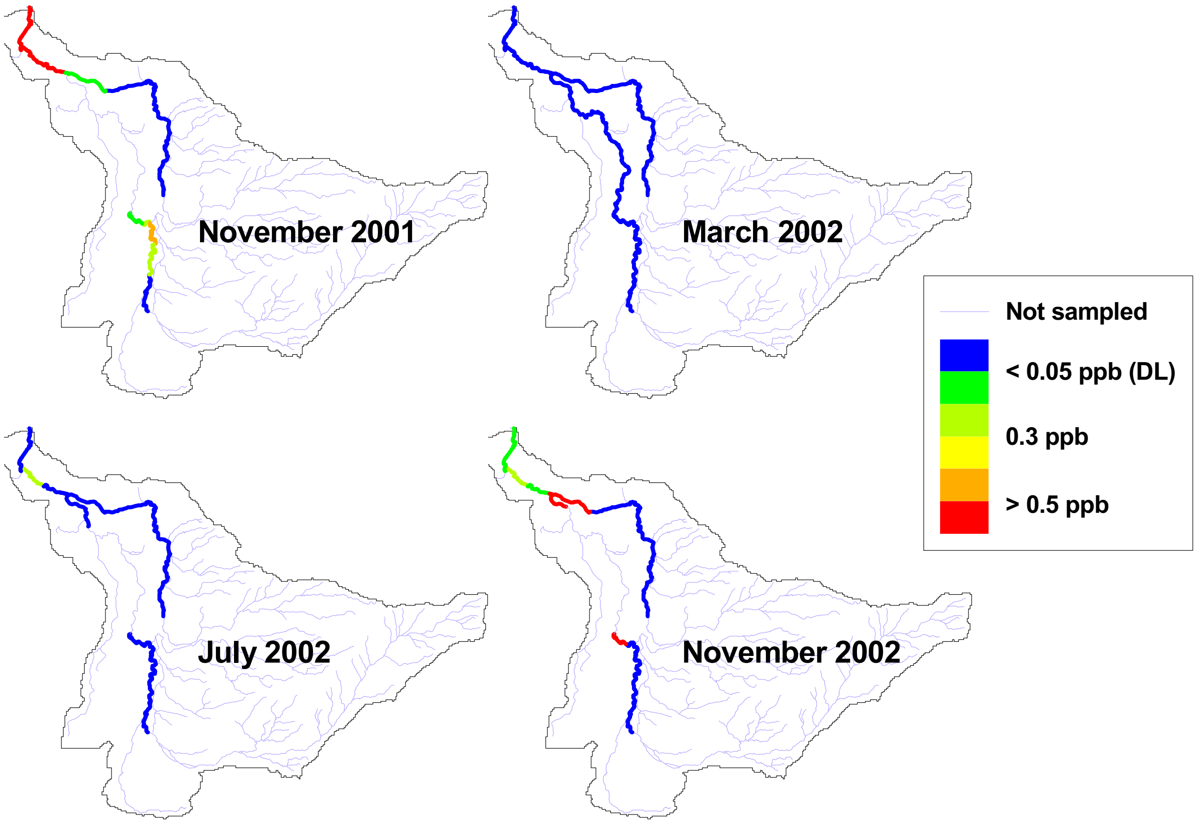


Figure 3.30. Monitored Imazalil concentrations

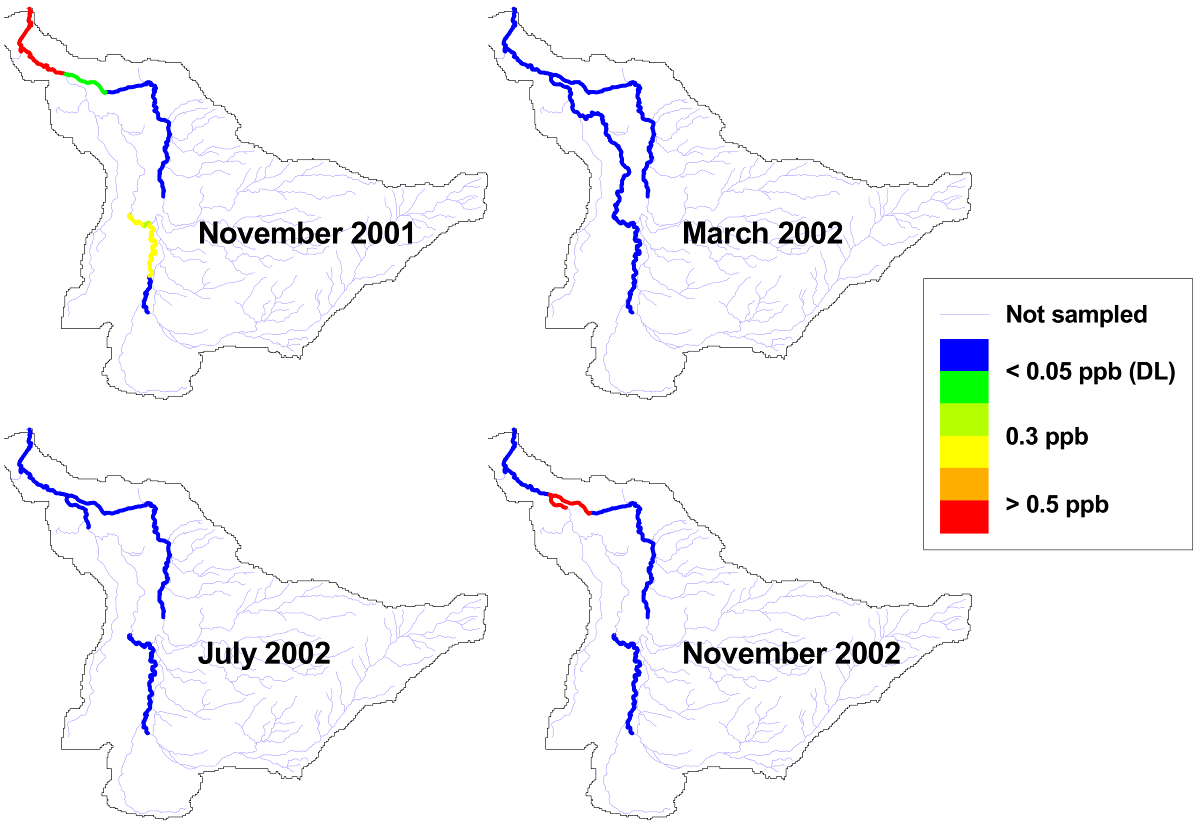


Figure 3.31. Monitored Thiabendazole concentrations

The pesticide analysis performed by the laboratory could detect other pesticides than the requested ones. Organophosphorus and Organochlorine pesticides were also detected by the chromatography analysis. Those detected pesticides represent mostly insecticides, which seem to be used in the majority of the basin's farms throughout the year, as shown in figure 3.32. In the figure, the total amount of detected pesticide concentration in every sampled stretch is displayed.

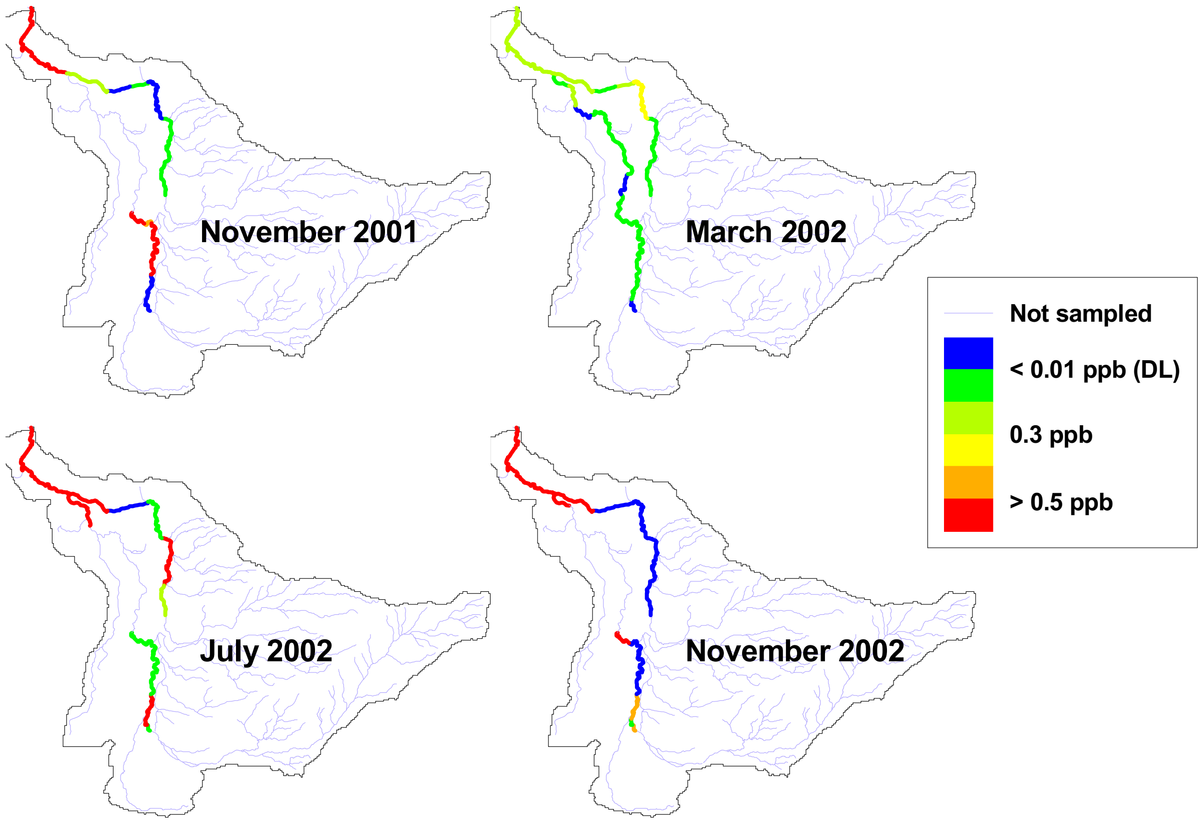


Figure 3.32. Total monitored pesticide concentrations

In the previous figure, maximum detected values were around 6 ppb (July and November 2002), which is significantly below the reported toxicity values for aquatic organisms. Based on that, the pesticide impact on the Chaguana basin on the aquatic biota is relatively low. However, detected pesticide values are exceeding the European maximum residue levels in water for human consumption (0.5 ppb for total amount of pesticides, and 0.1 ppb for one pesticide). Therefore, human health must be the main concern related to the pesticide usage in the Chaguana Basin.

3.12. CONCLUDING REMARKS

The data necessary to run a model sometimes is overwhelming, and are usually not so easily available. This chapter has gone through different methodologies to obtain reliable data to assess environmental problems in a case study: the Chaguana river basin. During the data evaluation, it was found that the existing state of the art in Ecuador regarding data housekeeping is still low. Although the procedures presented here can help to overcome that problem in other "data-poor" basins, they can not replace field measurements because uncertainties can be introduced in the model, and outcomes can differ from reality. All collected data has been organized and structured within a Geographical Information System. Thus, it can be available for future research or the basin stakeholders.

Another important finding in this chapter is the determination of the environmental baseline of the Chaguana river basin. There was no information regarding the environmental quality of the basin before this research began. Based on the sampling campaigns, it can be concluded that the basin is still not heavily polluted. However, data presented here should only be considered as a starting point to begin a monitoring programme on a frequent basis to develop a reliable historical data of several environmental parameters.

In the next chapter, a detailed description of the model evaluation is given. Selected models were executed with all data presented in Chapter 3. GIS was used to input some data into the models.