한국환경과학회지 제5권(제2호), 103~112, 1996
J. of the Korean Environmental Sciences Society

Sampling Study on Environmental Observations: Precipitation, Soil Moisture and Land Cover Information

Chul-Sang Yoo
Division of Environmental, Ocean and Water resources Engineering
Department of civil Engineering Texas A & M University
Colleges Station, TX 77843-8183, USA
(Manuscript received 15 January 1996)

Observational date is integral in our understanding of present climate, its natural variability and any change due to anthropogenic effects. This study incorporates a brief overview of sampling requirements using data from the First ISLSCP Field Experiment (FIFE) in 1987, which was a multi-disciplinary field experiment over a 15 km grid in Konza Prairie, USA. Sampling strategies were designed for precipitation and soil moisture measurements and also detecting land cover type. It was concludes that up to 8 raingages would be needed for valuable precipitation measurements covering the whole FIFE catchment, but only one soil moisture station. Results show that as new gages or station are added to the catchment then the sampling error is reduced, but the improvement in error performance is less as the number of gages or stations increases. Sampling from remotely sensed instruments shows different results. It can be seen that the sampling error at larger resolution sizes are small due to competing error contribution from both commission and omission error.

Key words : Sampling, sampling error, precipitation, soil moisture, land cover

1. Introduction

Climates studies hope not only to simulate present climate, but to understand the natural variability and to detect changes due to anthropogenic effects. Understanding the controlling factors and feedbacks is an important part of this process and particularly when predicting and assessing the impact of regional and global change (at least as far as the 21st Century). An integral part of this are observations required for:

1. data assimilation procedures to model
2. providing model boundary conditions
3. initializing model studies
4. model verification
5. conducting process studies

It is important that observations are readily available and objectively from both remotely-sensed and *in situ* data. In practice the available data are a combination of these two, often brought together the use of a physically-based conceptual model.

The purpose of this study is to consider a range of sampling techniques for use with either in situ or remotely-sensed data. Techniques to combine these two types of observational data are beyond the scope of this study. All observations contain an error of some magnitude and one hopes that by choosing the most appropriate sampling technique for any application, then this error can be minimized. Precipitation, land use, and soil moisture have been chosen for this study, with a view to examining the sensitivity to the sampling techniques discussed above.

2. First ISLSCP Field Experiment (FIFE)

This was a land-surface-atmosphere experiment based on a 15 km×15 km ares of predominantly grassland on the Konza Prairie, near Manhattan, Kansas. There was extended monitoring of satellite, meteorological, biophysical and hydrological data during periods in 1987 and 1989. The sampling stations were distributed over the catchment. Sellers et al. (1992) gives an overview of the experiment.

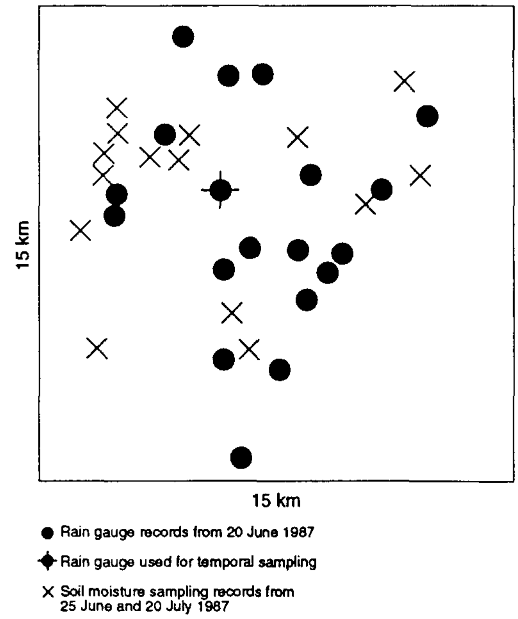

Fig. 1. Location of rain gauge and soil moisture sampling stations over the FIFE catchment, used in this study.

3. Sampling Study of Precipitation Data

In this analysis 21 raingages within the FIFE catchment were used. These collected data at 30 minute intervals, at locations shown in Figure 1. Two types of sampling analysis were performed: spatial and temporal A short rainfall event from 20 June 1987, Julian day 171, was selected for the spatial sampling analysis. This was only 3 hours long so limited unnecessary calculations. The following assumptions need to be made so that the spatial sampling errors from the raingages can be estimated:

1. the averaged rainfall information over the given area represents the 'true' rainfall
2. the area considered is small enough to disregard the effect of storm movement in time
3. the observation timing error among gages is negligible.

The Root Mean Square error (RMS error, E) between the observed rainfall value at one point within a region and the 'true' rainfall for that region is calculated at each time of observation (that is every 30 minutes) using the following expression:

$$E=\Sigma\sqrt{(OBS-TRUE)^2}$$

where OBS is the observed value and $TRUE$ is the 'true' value of rainfall. The catchment can therefor be divided into n regions and the RMS error for n gages calculated, leading to an estimate of the least number of gages required to observe the rainfall event to the required accuracy. Figure 2a relates the total rainfall recorded at each gage to the sampling error. This shows that approximately 4 gages are necessary to observe the precipitation to within 10% of the mean. Figure 2b shows rainfall sampling a each time within the storm event integrated over the duration of the storm.

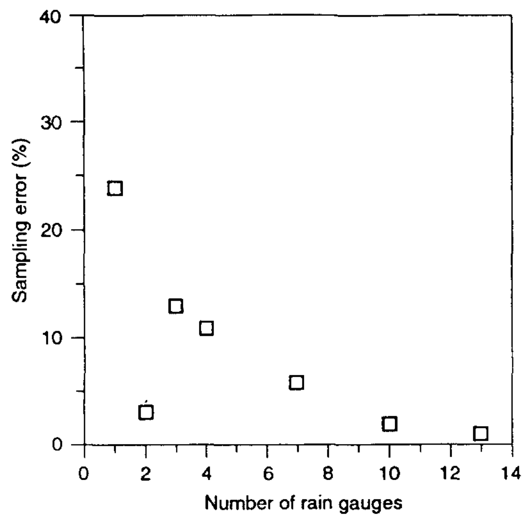

Fig. 2a. Sampling error as a function of the number gauges, for total rainfall over the event (without including temporal variability).

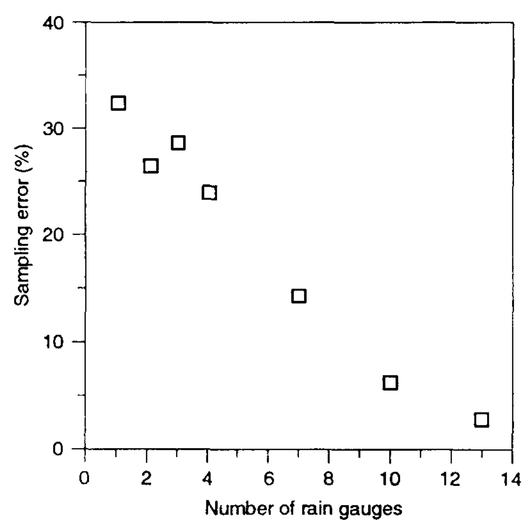

Fig. 2b. Sampling error as a function of the number gauges, for rainfall at 30-minute intervals intergated over the duration of the event.

This accounts for the temporal variability within the event not considered in Figure 2a. in this case sampling errors are greater and 6 to 8 gages are required to achieve a 10% error.

The temporal sampling analysis examined a

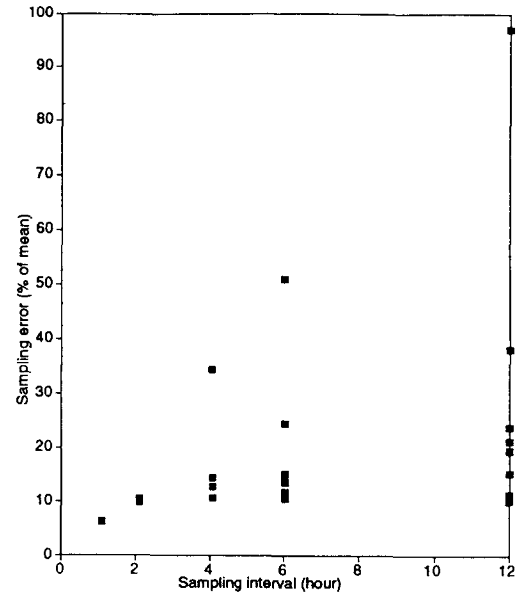

Fig. 3. The increase in error as a function of both sampling interval and time of sampling within that interval (shown by the vertical variability at each sampling interval)

time series of data from midnight on 24 June (day 175) to midnight on 9 July (day 190), for one gauge selected at random from those available. The sampling errors introduced due to the intermittent sampling of the rainfall were estimated using a similar technique. It was assumed that methodological condition recorded at each time (that is wet or dry) continues until the next sampling time. The data shows highly varying rainfall intensities, that introduces significant errors into the estimation. The sampling intervals were increased and the errors examined as shown in Figure 3. It can be seen that, as expected, the sampling error increases as the sampling interval increases. Figure 3 further shows the variability of error due to the start time of the integration within the sampling interval, demonstrated by the spread of points at each sampling interval. Figure 4 shows the variability of results from selecting different start times to the

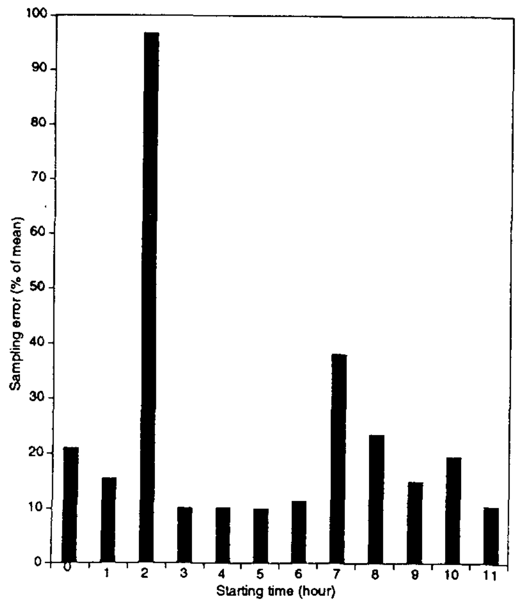

Fig. 4. An example of variation in error given different start times over a 12-hour sampling interval.

sampling interval during a 12 hour period. One large duration precipitation event during the observing period could affect sampling error significantly, especially when the observing period is relatively short such as in this case. If a longer data set had been used with a finer temporal resolution, more reliable results would have been obtained.

4. Sampling Study of Soil Moisture Data

Soil moisture measurement are required to improve estimations made using process scale studies. They are also used for model validation. The measurement accuracy required depends on the model type, and purpose. In this study, a representative soil moisture value has been calculated for the FIFE catchment.

The hypothesis for this part of the study is that soil moisture varies over the area of the catchment, but can be represented by some point measurements, using the assumption that the area can be considered homogeneous. The average soil moisure calculated using all of the stations over the catchment was assumed to be the 'true' soil moisture value.

In situ gravimetric soil samples were taken over the first 10 cm of the soil profile at 15 stations across the FIFE catchment. As can be seen (Figure 1) these sampling stations were not evenly distributed over the experimental area. An airborne radiometer detecting the top 5 cm soil water content, the PushBroom Microwave Radiometer (Wang et al., 1992), PBMR, was flown over parts of the catchment at periods throughout the experiment.

Data was selected for two days: the 25 June 1987, Julian 176, collected barely hours after a heavy rainfall event of 26 mm. The PBMR data was also flown on that day. The 20 July, day 201, was the second day selected, which was 15 days after a rainfall event, when the soil was much drier.

A preliminary study was undertaken within the data: all of the samples at each station were averaged and the mean values compared to that of the whole catchment (Figure 5a and 5b). It can be seen that the individual mean values at each station vary considerably about the catchment mean. Therefore the soil moisture field can be seen to have a natural variability.

The number of sampling stations was then reduced and the mean values of the remaining sites calculated. This was repeated several times and the results shown in Figure 6a and 6b.

It can be seen that the errors introduced by reducing the number of stations are small, and thus it can be concluded that a few sites can give a representative soil moisture value. While the % error is small however many stations are selected the error variability is reduced by having more sta-

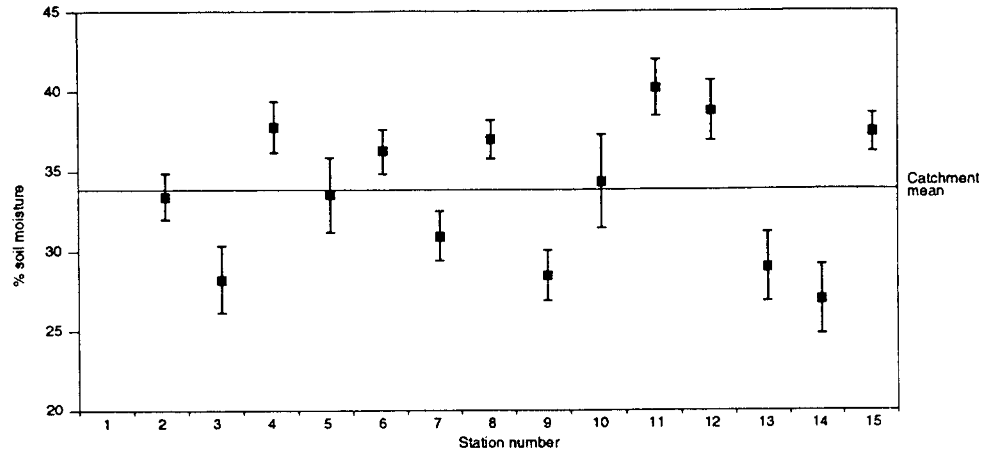

Fig. 5a. Mean and standard deviation of soil moisture for each sample station on day 176.

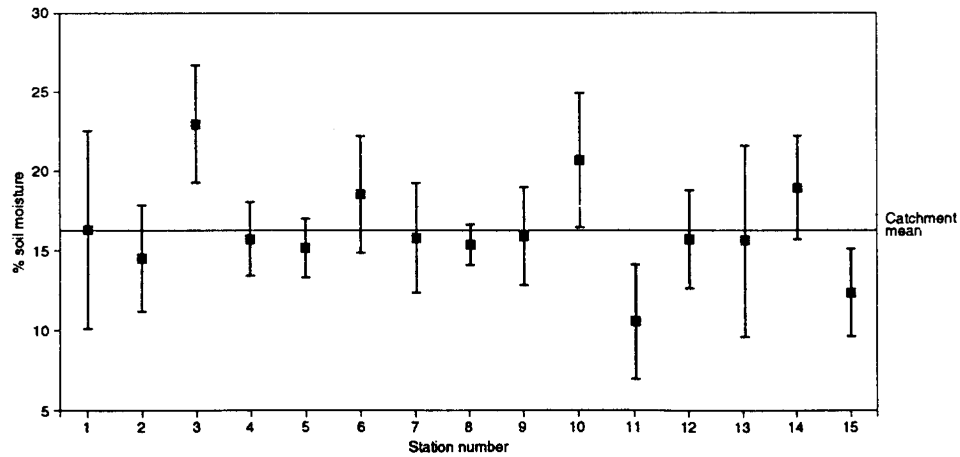

Fig. 5b. Mean and standard deviation of soil moisture for each sample station on day 201.

tions in the sample. To examine further the number of stations required, a study of sampling strategy was carried out. The mean value of an area of the catchment was compared to the value of a central point within the area. The catchment was then divided into segments of equal areas, and the error between the mean value for each segment and the representative point was considered for different number or segments. It was found that the errors did not decrease noticeably, as the segment size decreased. There it can be concluded that, in the case of soil moisture needs only to be measured in one or two places over the catchment. It must noted that there are extreme cases, where the soil moisture appears anomalous, but these cases can be reduced by taking a large number of samples at each station.

These analyses have been performed on the assumption that the area is homogeneous: in reality there are effects from the 5 different types of soil; different vegetation cover and topography. These variables all affect soil moisture, resulting in variability over the whole catchment. This can also be seen from the PBMR image shown in Figure 7.

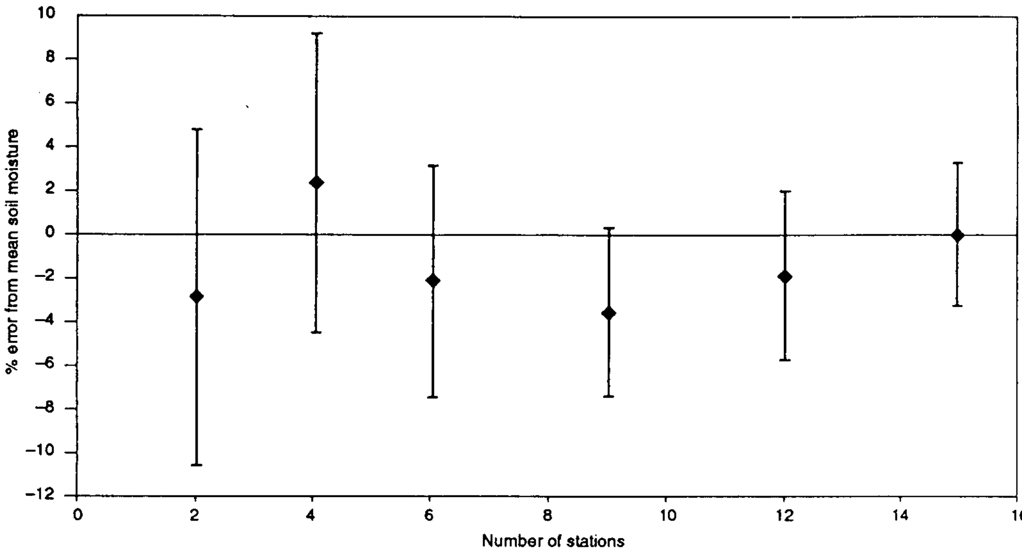

Fig. 6a. The % error between the catchment mean and the mean obtained from an increasing amount of stations, for day 176.

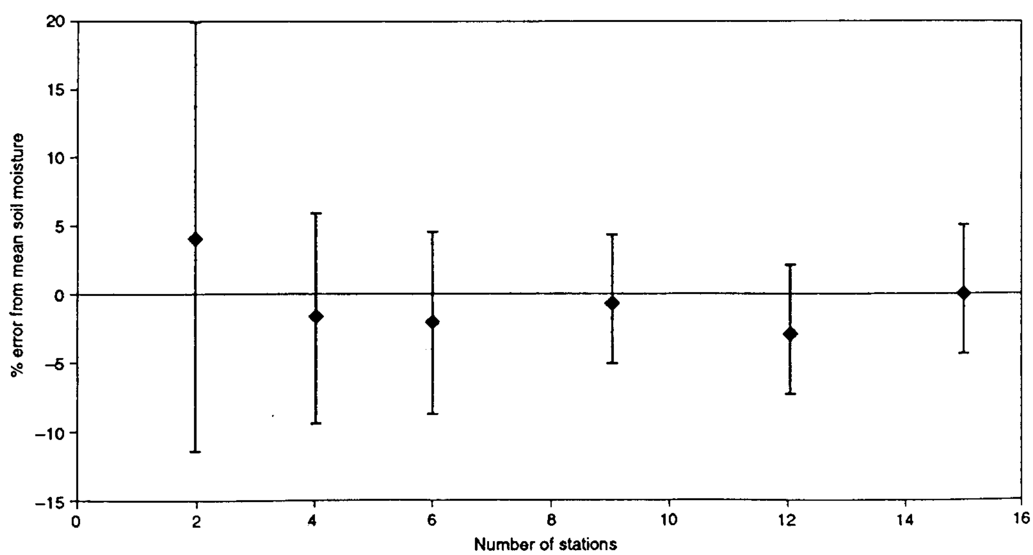

Fig. 6b. The % error between the catchment mean and the mean obtained from an increasing amount of stations, for day 201.

The analysis on the other hand show that when the area is sampled by several different methods the mean soil moisture agrees with the 'true' value to withir. standard deviation or 10%, and can be considered as a random variable. The required accuracy of the soil moisture measurements depends on the needs of the user, but 10% error is taken to be acceptable case. This result holds for both day 176, when the soil is wet, and day 201 when it is drier.

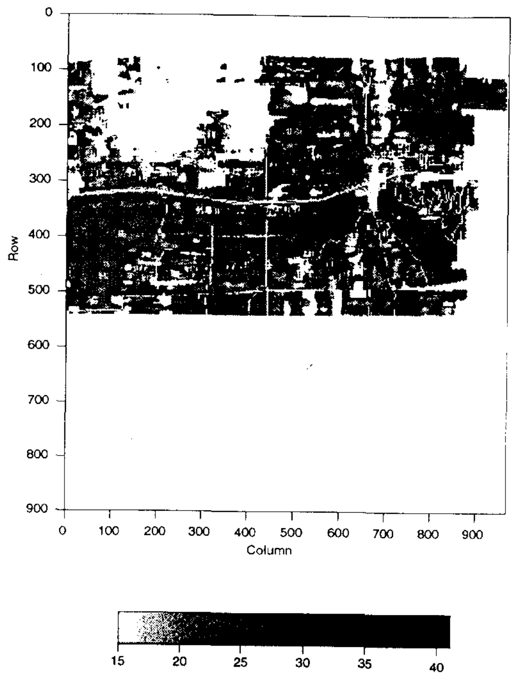

Fig. 7. PBMR measurements of % soil moisture over the FIFE catchment, on day 176.

5. Sampling Study of Land Cover Inforamation

The FIFE catchment was largely grassland, but with about 5% trees in the bottom of the river valleys. The wooded area is organized in narrow lines along these valleys. In order to observe the correct proportion of trees it is necessary to be able to resolve these narrow features. This study examines the minimum resolution required in order to estimate the proportion of trees currently, using SPOT 30 m data for the catchment. These data (Strebel et al., 1992a, b, c) at full resolution gives a close estimate of the proportion of trees. One band of the data (band 2) was used for simplicity, and a reflection band was selected, with two thresholds, to estimate the proportion of trees. Bi-directional reflectance distribution functions and shading effects on the hill slopes meant that grass could be either brighter, or darker than the area classified as trees, and so for future studies it would be more effective to use the full multi-spectral information.

From first principles, a feature has to be resolved by sampling at least twice its scale length to be estimated without aliasing errors, otherwise it is subject to commission error. At larger scales, not identifying areas of woodland which emit result in omission errors. Both types of error are evident when sampling at the range of pixel resolutions.

The original 30 m data were degraded by averaging into larger pixels, and the mean reflectance calculated for those larger pixels. The data were then threshold for each of these larger with the same threshold values for reflectance to estimate the proportion of woodland pixels. The original image is shown in Figure 8a and the degraded image in Figure 8b (at a pixel size of 120 m) and 8c (at a pixel size of 480 m). All adjacency and non-linear effects were assumed negligible for this problem. A more complete study would have had to take into account instrumental effects along with other effects on the surface radiance, such as the atmospheric correction. These results therefore only illustrate the principle observations that may be expected from a larger study.

As already mentioned sampling from remote instruments suffer errors due to both commissioning: the effect of aliasing due to insufficient sample length and omission: the sampling interval is too large to resolve the features. It was assumed that the % of trees resolved at 30 m resolution is the true woodland cover: of the order of 5%. Figure 9 shows that larger commission errors, even for the 60 m pixel sizes. This analysis probably over estimates these errors, but they are not negligible. It may be seen from the data that much of the woodland is in very narrow strips, and a doubling

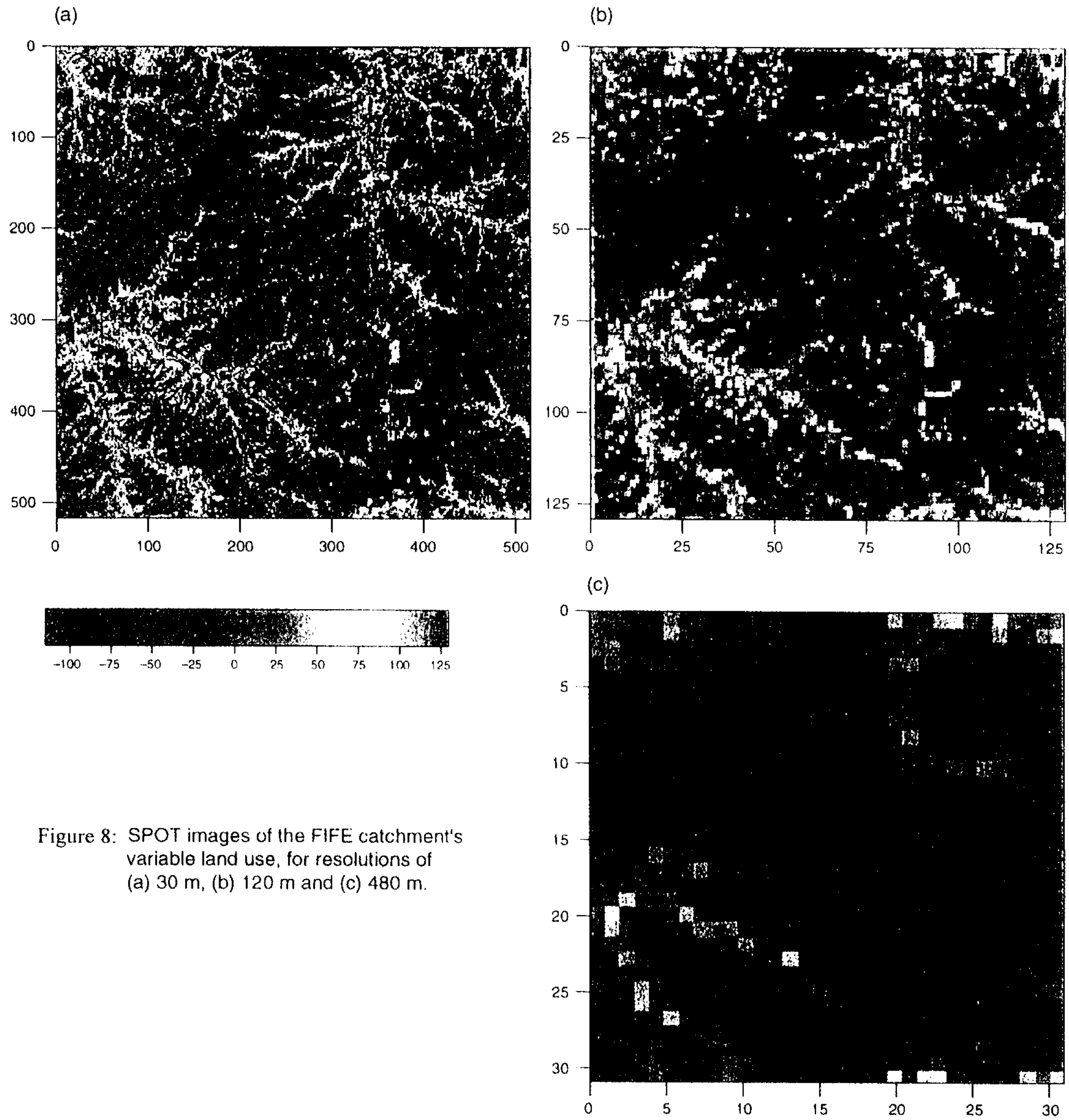

Fig. 8. SPOT images of the FIFE catchment's variable land use for resolutions of (a) 30m, (b) 120m, (c) 480m.

of the pixel size will indeed result in larger errors.

At larger pixel sizes, greater than 300 m resolution, omission errors dominate. At just over 1 km pixel size, the two errors approximately balance. Data from a satellite viewer such as AVHRR should therefore give errors in the estimation of woodland areas on the FIFE catchment that are smaller than viewers with half resolution, but still not as good as those that resolve the woodland currently.

Data such as AVHRR has sampling errors as low as instruments with half the resolution because increasing errors due to omission and commission act against each other. The spatial distribution of

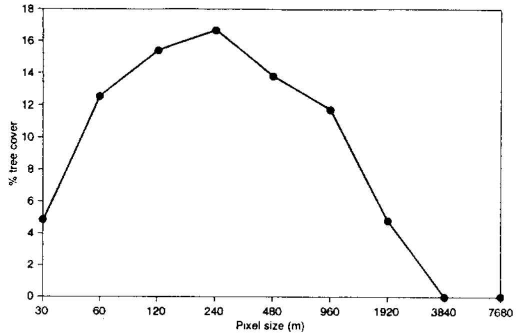

Fig. 9. The % tree cover over the FIFE catchment as a function of sample pixel size. The value of 5% at 30 meters resolution is assumed to be the catchment cover.

a land use type can also affect its ability to be sampled. Figure 10 shows as example of this: as the length of the woodland area increases the classification of the continuity of the woodland features becomes harder. This demonstrate the difficulty in sampling land use features, that are not evenly distributed in space.

6. Conclusion

This study examined a number of sampling methods from both in situ and remotely-sensed data. In particular the study concentrated on sampling of precipitation, soil moisture and woodland cover over the FIFE catchment. The FIFE data was assumed to be the 'truth' and variations of the sampled data provide an estimate of deviations from this 'truth'.

In the precipitation study it was noted increasing gauges in the catchment improves the sampling of rainfall. In this study around 4 gauges were required to reduce the error to 10%, by considering the total rainfall over an event. If the temporal variability of the rainfall was included the number of gauges increased to between 6 to 8. It was also found that the % error was affected by the sampling interval and collection time within the interval. The sampling of soil moisture over the study catchment was found to be much simpler. Insensitivity in sampling error to measurement stations meant that a small number are required to provide representative values for the catchment as a whole.

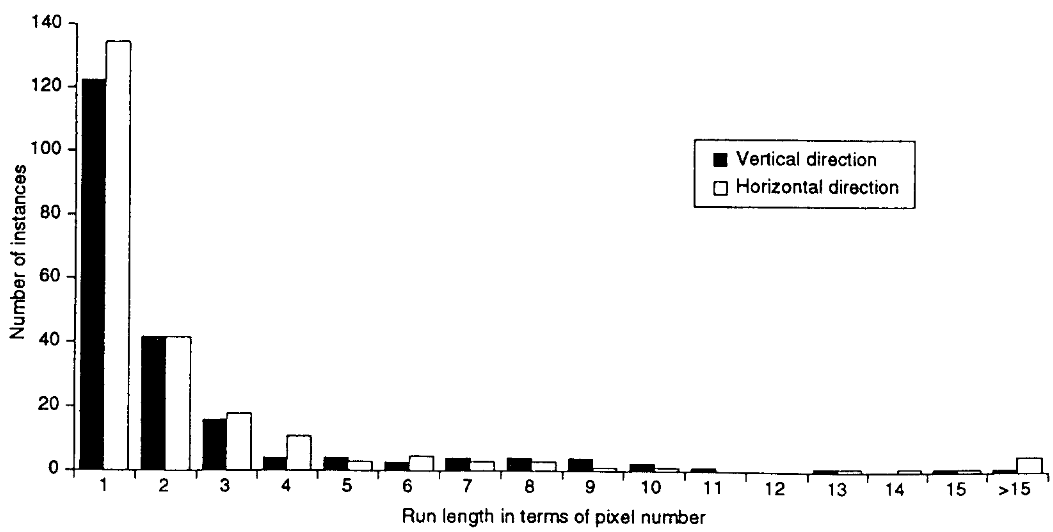

Fig. 10. The identification of woodland features across the FIFE catchment as a function of run length.

In sampling woodland cover it was found that degrading the sampling resolution did not necessarily increase classification error. In fact errors due to commission and omission of the woodland region acted against each other to reduce classification errors at larger resolution. This also appeared to be a function of the feature's continuity.

Although many of the results discussed here are based on a small data set, the methods are applicable to larger studies. The study aims to suggest some important areas of sampling strategy, particularly in relation to obtaining data for use with process models and larger scale Global Climate Models (GCMs).

Acknowledgments

This work was carried out as apart of a EC sponsored Summer School for Environmental Dynamics, at the Institute Veneto, Venice. The support of the technical staff at the institute is gratefully acknowledged.

References

Sellers, P.J., F.G. Hall, G. Astar, D.E. strebel and R.E. Murphy, 1992. Am Overview of the First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment (FIFE), *J. Geophys. Res*, 97, 18345~18372

Strebel, D.E., D. Landis, J.A. Newcomer, J. DeCampo and B.W. Meeson, 1992a. Collected Data of the First ISLSCP Field Experiment, prototype: satellite, aircraft and ground measurements, May- December 1987. Published on CD-ROM by NASA.

Strebel, D.E., D. Landis, J.A. Newcomer, S.J.Goetz and B.W.Meeson, P.A. Agbu and J.M.P. McManus, 1992b. Collected Data of the First ISLSCP Field Experiment, volume 4: ASAS and PBMR Imagery 1987 and 1989. Published on CD-ROM by NASA.

Strebel, D.E., D. Landis, J.A. Newcomer, S.J. Goetz, B.W.Meeson,P.A. Agbu and J.M.P.McManus, 1992c. Collected Data of the First ISLSCP Field Experiment, volume 5: Image derived from satellite, aircraft and geographic date, Published on CD-ROM by NASA.

Wang, J.R. S.P. Gogineni and J. Ampe, 1992. Active and Passive Microwave Measurements of Soil Moisture in FIFE, J, Geophys. Res., 97, 18979~18986.