LANDSAT-TM DATA FOR MUNICIPAL ENVIRONMENTAL PLANNING ? STUDIES OF VEGETATION INDICES IN THE URBAN AREA

Matthias Achen
Assistent Professor, Department of Geography
University of Heidelberg, Germany
ISPRS-Commission VII

ABSTRACT

The aim of this study is to determine the accuracy with which the percentage of vegetation cover (PVC) in urban areas, can be recorded by digitally processing Landsat-TM data. The area analyzed consists of the settled area of Heidelberg, Germany; the reference data were yielded from the analysis of aerial photography. Various vegetation indices (NDVI, PVI, SAVI, PC2) were determined using geocoded TM data. The correlation with the PVC referent ($n=12,585$) was calculated for every index on the basis of 50-m grid squares on a scala calibrated from 0 to 100; they are found on the level of $r=0.92$. Comparative advantages are demonstrated by PVI in the lower PVC range and by NDVI in the upper PVC range. SAVI and NDVI display relatively similar results for the area analyzed; masking the settled area prior to the principal component transformation is favorable for PC2. Residues are largely caused by the dependency of all indices on the composition of surface materials and on shadow formation and by radiometric and geometric errors in the TM data. Moreover, the vegetation cover partially changed between the time of the satellite and the aerial pictures; this is why the results for all indices were in reality (even) more favorable. There is even a correlation in the range of $r=0.97/0.98$ between the indices and PVC for selected test areas made on the basis of 99 (building) quadrangles. The results of the study show that even on the municipal level, data from satellite photographs could already be used today for environmental planning.

Keywords: Landsat-TM, Image Processing, Vegetation Indices, Percentage of Vegetation Cover, Urban Area, Municipal Environmental Planning, Heidelberg (Germany)

1. INTRODUCTION

This article reviews the most important results of an empirical investigation of the accuracy with which the proportion of the total surface area covered by vegetation, i.e., the percentage of vegetation cover (PVC), can measured by digitally processing Landsat-TM data with the help of vegetation indices (ACHEN 1992). The settled area of Heidelberg is analyzed, an area displaying the basic types of urban construction and land use. Visual analysis of color infrared aerial photographs provided the reference data.

The state of technical data and information on the environmental situation in urban areas requires decisive improvement; this forms the general backdrop to the study. At least in the Federal Republic of Germany, official statistics in this sector do not yet include relevant, spatially related data. However, since such data is indispensable to the increasingly important political concept of "ecologically oriented urban development", a suitable procedure for measuring so-called environmental indicators must be found. Procedures are evaluated in terms of expense, time required, relevance, accuracy, and degree of differentiation of data, as well as the resolution and precision of geometric information (RADERMACHER 1989, 19).

Basically, there is no procedure that optimally satisfies these requirements. In data collected on the ground, differentiated indicators (such as the amount of vegetation) can be very accurately measured; however, the cost in time and money preclude the extensive use of this procedure. Visual evaluation of (color infrared) aerial photographs (depending upon their scale) must be judged as less accurate and differentiated, and its costs in time and money are only somewhat lower; this is presently the procedure most frequently used for collecting environmental data at the municipal level. Though the digital processing of satellite photograph data must be evaluated as at best satisfactory in resolution, precision, and the differentiatedness of data, its expenditures of time and money are considerably lower. Detailed data have yet to be published on the accuracy of such studies. Generally, one must seek an "ends-means-oriented procedure" (HÜBLER 1986, 465) for data collection. The aforementioned criteria have to be weighed against each other in a procedure where, given limited government financial resources, the importance of "expenditure of time and money" as a criterion is not to be underestimated.

The urban area cannot yet be counted among the objects of research for which digital satellite remote sensing has become a significant source of information. This may be explained in part by the incongruence between land use and land cover and the high percentage of mixed pixels provided by currently available sensory systems. Thus, the author is only aware of two publications on vegetation indices in urban areas: the study by KERL (1989) using Landsat-TM data on Munich, in which, however, the statistical evidence is incomplete; and the study by FORSTER (1983/1985a) using Landsat-MSS data on Sydney, in which the results are only suitable in part for operationalization in environmental planning.

This study thus aims to study the quality of various vegetation indices in measuring the percentage of vegetation cover (PVC) as an parameter relevant to urban ecology (KENNEWEG 1975) and to carry out this study on two levels: for 50-m grid squares covering the entire settled area of Heidelberg (31.5 km^2) and for (building) quadrangles from selected test areas in Heidelberg (1.16 km^2). The empirical studies have been supported by the German Institute for Aeronautical and Space Research (Deutsche Forschungsanstalt für Luft- und Raumfahrt; Oberpfaffenhofen) and the city of Heidelberg.

2. AREA INVESTIGATED

The area analyzed consists of the settled area of Heidelberg, defined here on the basis of the land use plan. It is primarily made up of residential, commercial, mixed, and public building areas. Furthermore, areas categorized as special building areas, areas of public utilities and waste management plants, public green spaces, allotments, cemeteries, traffic areas, and water areas were considered part of the settled area if they were largely enclosed by areas from the first set of categories. The same holds for small areas of agricultural use. On the other hand, smaller-sized settled areas located a clear distance from other settled areas were excluded. In this way, the settled area of Heidelberg corresponds roughly to a "restricted area" (geschlossene Ortschaft) and represents an area of about 31.5 km^2.

3. REFERENCE DATA

Color infrared slides (Kodak Aerochrome Infrared 2443) from a flight over the boundaries of Heidelberg formed the basis for the aerial evaluation. The flight, commissioned by the city of Heidelberg was carried out by Photogrammetrie GmbH (Munich). The pictures were taken on 17 August 1987 around 12:00 pm from a height of about 1500 m; their scale was on average ca. 1:5000. Such aerial photographs allow one to measure PVC highly accurately on the 50-m grid level on a scale from 0 to 100.

In order to measure vegetation areas, the slides were enlarged to a scale of 1:2000 and partially fit into land registry maps of the same scale. Each of the 50-m x 50-m areas to be evaluated was subdivided into 100 5-m x 5-m areas by means of a grid foil. Those grids covered with vegetation were then counted, which, given the size of the area analyzed, entailed a considerable expenditure of time and energy. Errors in the evaluation of the aerial photography mainly result from adjustment problems due to insufficient ground control points (GCP), variable photographic scale due to radial deformation, and measurement problems given very heterogeneous vegetation structures and regions of shadow. Control studies in selected test areas indicate that the accuracy of the PVC data, aside from a few exceptions, is overall within ca. five percentage points.

Areas of agricultural use were always classified as vegetation areas, regardless of the actual conditions on the date of aerial filming. This measurement decision resulted from the following circumstances: The Landsat-TM data were from the second half of June, since this is a time of less shadow formation. It is also a time when areas of agricultural use are largely covered with vegetation. However, it is impossible to reconstruct the vegetation structure of such areas at the end of June from aerial photographs that were taken in the middle of August.

The arithmetical mean of PVC is 48% for the area analyzed. The settled area of Heidelberg is thus almost half "green." Compared to other cities in ecological terms, this has to be classified as a favorable result. Within the city, there are great differences in vegetation structure, reflected on the administrative level in the arithmetic mean of the 33 municipal districts. Some of the districts defined by a high degree of commercial use present values under 25%; in contrast, districts largely marked by low-density residential development usually show values of over 60%.

4. EMPIRICAL RESULTS

4.1 Data Base

The basis for the calculation of the vegetation indices was provided by the Landsat-5-TM scene 195/26 (WRS) of 27 June 1986, which was received via relay satellite (TDRSS) and systematically restored by EOSAT. Heidelberg is located in the upper third of the satellite's receiving range and is not marked by the influence of clouds, haze, or other disruptive atmospheric factors. Even though another TM scene 195/26 was also available, which was received by the ESA ground station in Fucino, Italy on 17 August 1987 (i.e., at the same time as the aerial photograph referent) and systematically restored by ESRIN, the EOSAT scene provided the focus for the empirical studies performed here. This is due to two decisive comparative disadvantages of the ESA scene, which are not offset by the temporal parallelism of the satellite and aerial data:

- Inadequate systematic restoration by ESRIN
- Higher percentage of shadows at the center and periphery of the area analyzed

4.1.1 Systematic Restoration of TM Data It is generally necessary to geometrically restore the raw data of a satellite-conveyed scanner system, since neither the movement of the satellite nor the filming process function without disturbances. The quality of the systematic restoration performed by ESRIN, however, has been (as yet) inadequate. The uncorrected ESA data are marked by two elementary kinds of errors, proof of which can be directly furnished both optically and by computation:

- About every 60 lines, a line is doubled.
- At irregular intervals (but always in a multiple of 16, if the doubled lines are not counted), a line segment is moved out of place.

These two kinds of errors, which can also be found in other ESA scenes, probably have, according to EHRHARDT (1990), the following causes:

- During filming, the satellite flies too low and thus too rapidly over the area, creating scan gaps between the individual sweeps of ca. 8 m or 0.27 pixel width. These are compensated for in ESRIN by a doubling of ca. every sixtieth line, in order to globally optimize the geometry.
- The earth's rotation and the fluctuations in mirror rotation are incorrectly rectified in ESRIN, leading to repeated line segment misalignment in both directions.

There is in part appreciable disruption of the geometry of the ESA data, occurring locally (i.e., on the pixel level); thus, location errors of over 90 m (!) within the area analyzed occur at the uncorrected level. It is possible to eliminate the doubled lines and to realign misplaced line segments (at least with an accuracy of ± ca. 15 m) prior to correction; nevertheless, it is more reasonable for the evaluation of urban-area data to use TM scenes that (for example) have been systemically restored in EOSAT.

4.1.2 Shadow Formation in TM Scenes The standard studies of the features of vegetation indices point out that such indices can sometimes vary greatly due to the angle of sunshine (i.e., the size and type of shaded areas) (WARDLEY 1984, USBECK 1989). In the area analyzed, two different kinds of regions of shadow are found, namely, on the periphery of the area analyzed, in the TM pixel sector, caused by relief (e.g., in the Neckar valley); and at the center, in the TM subpixel sector, caused by individual objects (above all, houses and trees).

For this reason one can expect that the measurement of PVC by means of a vegetation index may be strongly impaired by shadows as a special spectral signature. This effect is minimal for scenes that are filmed when the sun is at its highest position, which applies to the EOSAT scene of 27 June 1986. Let us take into account the time of local overflight of Landsat 5 (ca. 9:30 am Central European Time), since it is hardly likely that this time can be changed for central Europe. Given this time, shadow lengths in the area analyzed are reduced by at least about one quarter in the EOSAT scene as compared to the ESA scene; they nevertheless remain of considerable size.

4.2 Geocoding of the TM Subscene

In the evaluation of TM data from an urban area, usually marked by heterogeneous and small-sectioned land-use structures, it is very important to geocode highly accurately, since otherwise (depending on local conditions), large errors can result. Thus, for the present study, the EOSAT subscene has been corrected by mapping it onto the Gauss-Krüger coordinate system using a third-degree transformation function with a comparatively high number of GCPs (19). The GCPs are taken from the Deutsche Grundkarte (Basic German Map) 1:5000, which in contrast to other topographical maps is not generalized. Resampling was performed via cubic convolution, which reconstructs, with as a rule the smallest degree of error, the spectral signals for individual pixels in the corrected picture in comparison to the nearest neighborhood or to bilinear interpolation. Pixel size in the geocoded picture is 25 m, which ensures compatibility with the aerial referent after aggregation.

4.3 Vegetation Indices in Urban Areas

In previous years, a great number of vegetation indices for measuring PVC have been developed for regions primarily marked by agriculture and forestry (PERRY/LAUTENSCHLÄGER 1984). Most indices take advantage of the differences in reflection that different surfaces possess in the red and infrared sector of the electromagnetic spectrum. In the application of these procedures to an urban area, it must be noted that the reflective properties of sealed surfaces are very similar to those of ground free of vegetation; however, as a rule, the former does not display as pronounced an increase in the degree of reflection as one moves from the visible to the infrared sector. Vegetation indices are usually formed as linear combinations and/or as ratios from the reflection in the red and (nearby) infrared wave-length sector. In the present study the following indices were examined:

- Perpendicular Vegetation Index (PVI)
- Normalized Difference Vegetation Index (NDVI)
- Soil Adjusted Vegetation Index (SAVI)
- Second Principal Component (PC2)

4.3.1 Perpendicular Vegetation Index PVI was developed by RICHARDSON/WIEGAND (1977) following the tasseled-cap concept and is based on the existence of a soil-line in the scattergram from TM3/TM4. The greenline is found orthogonally to the soil-line; the PVI of a pixel is determined by projecting the digital number (DN) onto this line. In constructing the PVI, the DN combinations of at least two pixels are necessary in order to determine the slope of the soil-line. If α represents the angle between the soil-line and the TM3-axis, then: PVI = $\cos\alpha$ TM4 - $\sin\alpha$ TM3

4.3.2 Normalized Difference Vegetation Index NDVI also presupposes the existence of a soil-line in TM3/TM4. However, whereas in PVI DNs with the same index value are found on lines that run parallel to the soil-line, these isopleths in NDVI converge at the origin. No reference pixels are necessary for construction. The following holds: NDVI = (TM4 - TM3)/(TM4 + TM3)

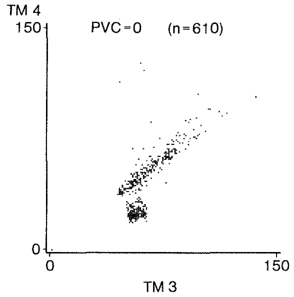
4.3.3 Soil Adjusted Vegetation Index SAVI was developed by HUETE (1988) as a "mediating" index between NDVI and PVI. Just like NDVI, SAVI possesses a convergence point of the vegetation isopleths; it is located, however, not at the origin, but at the angular bisections in the third quadrant. The distance between the convergence point and the origin has to be selected according to the situation and is determined above all by the PVC of the given area: the lower the PVC, the greater the distance. If L gives this distance in gray values (after multiplying it by $\sqrt{0.5}$), then: SAVI = (TM4 - TM3)(1 + (L/100))/(TM4 +TM3 + L)

4.3.4 Second Principal Component According to the study by INGEBRITSEN/LYON (1985), PC2 can be interpreted as a vegetation index; however, as yet there is at most only one quantitative study concerning its quality as an index (FÖRSTER 1985a). No reference pixels are necessary for carrying out a principal component transformation.

4.3.5 Calculation of the Indices Calculating NDVI was unproblematic. For PVI, the PVC reference data were used. Here, the soil-line was formed as a straight line of "medium" regression (α=41.8) on the basis of aggregated TM3/TM4 DNs of all 50-m grid squares where PVC=0 (however, all grid squares with water, i.e., the Neckar river, were excluded). For these 295 points, the correlation is r=0.934 (cf. Fig. 1). PC2 was calculated on the basis of the geocoded subscene of Heidelberg (520 x 520 pixels) with a geometric resolution of 25 m: either with all pixels from the subscene ($PC2_T$), or after masking the subscene, only with the pixels from the settled area ($PC2_M$). The second version is based upon the supposition that within the principal component transformation, spectral information from wooded or agriculturally utilized lands interferes with the quality of PC2 as a vegetation index. The construction of SAVI is presented in section 4.4.2.

4.3.6 Gray Value Analysis of Selected Grid Squares The reference data enable one to examine the (digital) gray values in TM3/TM4 not only for grid squares without vegetation or water, but generally for squares with any given PVC whatsoever. Figure 1 contains the two-dimensional gray values in TM3/TM4 where PVC=0, PVC=33(±1), PVC=66(±1), and PVC=100. In each of these cases, the digital values are arranged in clusters with different shapes. Each of the clusters displays a central tendency (i.e., longitudinal alignment), the slope of which increases nonlinearly from PVC=0 to PVC=100; the center moves counter-clockwise around the TM4-axis.

In none of the four situations presented are the DNs located on a straight line; especially in the middle PVC region, they display considerable variation, orthogonally to the central tendency. When PVC=100, most of the points are found relatively near the central tendency, though some values are a great distance away. There are primarily two main causes of such severe deviations: the dependency of gray values in TM3/TM4 on the angle of sunshine, which interferes with results, especially on slopes; and the method of measuring reference data for agriculturally utilized lands, whose spectral properties, if insufficiently covered by plants, correspond more to sealed surfaces than to areas of vegetation.



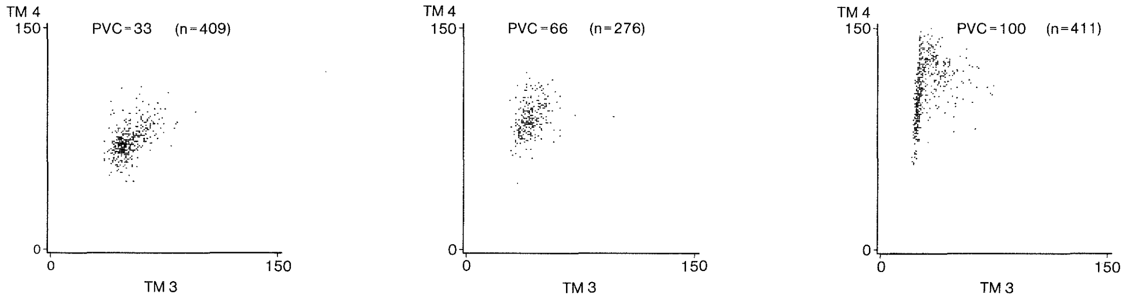


Figure 2: TM3-/TM4-gray values of grid squares with PVC=0 / PVC=33 / PVC=66 / PVC=100

The four situations indicate that the following three states of affairs hold for the area analyzed:

- The optimal convergence point of SAVI must not be far from the origin.
- In areas with a quantitatively low vegetation cover, PVC could be measured somewhat better by PVI than by NDVI; in areas with a quantitatively high vegetation cover, the opposite is true.
- Generally, this measurement probably includes considerable error, since the two-dimensional gray values of grid squares with a certain PVC in common are not located on a line, but are arranged in the form of a cluster.

4.4 Correlation: Uncalibrated Indices / Aerial Photographic Referent

4.4.1 Correlation: NDVI, PVI, PC2T, PC2M, and PVC In order to examine how accurately PVC is represented by the different vegetation indices, the correlation coefficient of every index in relationship to PVC was calculated. While it is indispensable for the operationalization of the indices to calibrate them by means of a transformation onto a scale from 0 to 100, this step is not absolutely necessary for a correlation analysis, since linear transformations do not effect the correlation coefficient. However, since the transformational function used in calibrating the indices in the present study possesses a nonlinear form, the correlation analysis of uncalibrated indices can "only" be of a preliminary character.

Table 1: Correlation prior to the calibration

	PVC	PVI	NDVI	$PC2_T$	$PC2_M$
PVC	-	.900	.900	.877	.911
PVI	.900	-	.970	.986	.984
NDVI	.900	.970	-	.940	.956
$PC2_T$	.877	.986	.940	-	.958
$PC2_M$	.911	.984	.956	.958	-

Table 1 contains the correlation coefficients of PVC with NDVI, PVI, $PC2_T$, and $PC2_M$ crosswise, prior to the calibration of the indices for 50-m grid squares. $PC2_M$ displays the most favorable result of all vegetation indices with a correlation of r=0.911 with PVC. NDVI and PVI take second place with r=0.900, followed by $PC2_T$ with r=0.877. The coefficient values provide a clear look at the quality of each of the vegetation indices in predicting PVC. The indices display a high level of agreement with each other for the area analyzed; for no combination does the correlation slip below r=0.935. The $PC2_M$ and $PC2_T$ values, which interestingly enough only display a (relatively) average correlation with one another at r=0.958, clearly show that the masking of the settled area prior to the principal component transformation has a positive effect on the quality of PC2 as a vegetation index. This result underscores the need to develop digital procedures for demarcating settled areas, something which OLBRICH (1990) has undertaken to do in a pilot study in terms of Heidelberg. Generally, PC2 is more highly correlated with PVI as with NDVI, which can be explained in terms of PC2's quality as a linear combination. Also striking is the fact that a remarkable correlation of r=0.970 exists between NDVI as a convergent isopleth model and PVI as a parallel isopleth model.

Figure 2 gives a representative look at the vegetation indices by depicting NDVI for the geocoded subscene of Heidelberg (13 x 13 km) with a geometrical resolution of 25 m. In this sector, NDVI ranges in value [-0.71, 0.77] and has been linearly transformed onto the (digital) gray-value scale [0,255]. The higher the NDVI, the lighter is the gray of the pixel. The correlation of r=0.900 between NDVI and PVC can be immediately recognized: surfaces densely covered with vegetation appear in light grays, well-sealed surfaces in dark grays. The different elements of the landscape contained in the subscene come into clear view in Fig. 2. In the east, the various shades of light gray of the Oden forest, largely covered with vegetation and displaying strong relief, can be seen. In the west, the wide array of gray shades of the Rhine plain can be made out, large sections of which are used agriculturally, but where extensive settled areas also exist. The parcellized structure of the agricultural lands can be clearly recognized in some parts, as well as the Neckar river, train lines, highways, and quite a few main streets, which appear as linear pictorial elements (free of vegetation). The densely built-up city center (the Old Town on the Neckar) and the old village centers of the incorporated parts of town are marked by dark shades of gray, just as the commercial areas, largely free of vegetation, in the western and southern parts of town.

4.4.2 Construction of SAVI The L value has to be set for this index. HUETE (1988) considers a value half the magnitude of the gray-value scale as reasonable for areas with a "medium" PVC, thus for TM data L=128. Since no previous studies of SAVI for an urban area are available, making it uncertain whether this value is applicable to the area analyzed (which displays a medium PVC of 47.9), L was varied over the range of five gray-value scales with an iteration amplitude of 16 gray-value gradations. The correlation reaches its maximum of r=0.909 precisely at L=128, confirming HUETE's results in a remarkable fashion; it then sinks uninterruptedly as L increases. However, the increase in the correlation between L=0, where it holds that SAVI=NDVI (r=0.900), and L=128 is not very great.

4.4.3 Correlation on the Level of Municipal Districts For the settled area of Heidelberg (as a whole area), the correlation between NDVI, SAVI, $PC2_M$, and PVI on the one hand, and PVC on the other, fluctuates more or less on one level. However, if the correlations are calculated for each of the 33 municipal districts, large differences between the individual districts arise for all of the indices. For example, NDVI ranges from r=0.969 to r=0.735. Moreover, notable differences between the individual indices also exist for quite a few municipal districts. Here, a certain regularity is recognizable, most clearly in the case of that pair of indices with the greatest differences (NDVI/PVI): PVI/PVC is more highly correlated than NDVI/PVC primarily for those districts whose PVC (average) is below the PVC (average) for the entire settled area (i.e., below 47.9); conversely, PVI/PVC is less highly correlated than NDVI/PVC mainly for those districts whose PVC is above the general PVC.

Table 2: PVC (average) and relationship between the correlation NDVI/PVC (r_1) and PVI/PVC (r_2) for the municipal districts

	$r_1<r_2$	$r_1>r_2$
PVC<47.9	15	2
PVC>47.9	4	12

Table 2 depicts the relationship between the two correlation coefficients and PVC for the 33 municipal districts in a 2 x 2 crosstable. The value of the corrected contingent coefficients is $CC_{corr}=0.762$ for this distribution; at the level of $\alpha=0.01$, the correlation is significant. This calculation can also be made in the same way for the five other pairs of indices; with the exception of the indexical pair SAVI/NDVI, the value of the corrected contigent coefficients was always above $CC_{corr}=0.5$, and the level of significance for the correlation over 0.95. While PVI displays comparative advantages in those municipal districts with a relatively low PVC (average) and NDVI comparative advantages in those districts with a relatively high PVC, $PC2_M$ and SAVI possesses their strengths in districts with medium PVC. In the direct comparison of $PC2_M$ and SAVI, $PC2_M$ tends to display advantages in municipal districts with below-average PVC and disadvantages in districts with above-average PVC.

Figure 1: NDVI in the geocoded subscene of Heidelberg (27 June 1987)

The differences between the individual correlations that arise in relationship to the given PVC found on the municipal district level also tend to find support on the level of the individual 50-m grid squares (n=12,585). In this context, NDVI displays a somewhat more advantageous structure in comparison to PVI in the upper PVC region and a somewhat less advantageous structure in the lower PVC region, which is in accordance with the findings of HUETE/POST/JACKSON (1985).

4.5 Calibration of the Indices

To operationalize vegetation indices it is necessary to calibrate them according to a scale [0,100] by means of a transformation, since otherwise only relative statements can be made. In the present study, this conversion was performed by means of a "quasi-cubic" function using the reference data. First, the arithmetical mean was calculated for those surfaces free of vegetation and water (x_1) and the mean was determined for the surfaces covered by vegetation (x_2). The latter possesses the following indispensable property for residual analyses: the average of the transformed index for the area analyzed ultimately corresponds exactly to the PVC average (47.9). Then, all indexical values $< x_1$ were set at 0, and all indexical values $> x_2$ were set at 100. Finally, all indexical values between x_1 and x_2 were linearly transformed into the interval [0,100]. This procedure is premised, however, on the existence of reference data for the entire area, something usually not given; thus, it needs to be modified for a general application of vegetation indices (outside of Heidelberg): x_1 is to be determined via representatively selected reference pixels of the various types of sealed surfaces and x_2 via corresponding pixels of the various types of vegetation.

4.6 Correlation: Calibrated Indices/Aerial Photography Referent

Table 3 contains the correlation coefficients of PVC and the transformed (T) indices, TNDVI, TPVI, TPC2M, and TSAVI crosswise for 50-m grid squares. The individual correlations between PVC and the various vegetation indices are all of approximately the same magnitude after calibration, reaching the level of r=0.92, which can be deemed a notable result. The correlation of TNDVI and TSAVI with PVC, at r=0.922/0.923, is slightly above TPC2M, at r=0.919, and TPVI, at r=0.915. The values show that the quality of all vegetation indices in predicting PVC can be clearly improved by means of a nonlinear calibration (cf. Table 1). By means of the quasi-cubic transformation, the agreement among the indices has also been clearly increased: the correlation for all combinations reached r=0.980. Since TNDVI and TSAVI are almost identical here at r=0.996, TSAVI will not be examined anymore in the following; the "greatest" difference exists between TNDVI and TPVI, with r=0.980.

Table 3: Correlation after the calibration

	PVC	TPVI	TNDVI	$TPC2_M$	TSAVI
PVC	-	.915	.923	.919	.922
TPVI	.915	-	.980	.989	.994
TNDVI	.923	.980	-	.984	.995
$TPC2_M$	.919	.989	.984	-	.992
TSAVI	.922	.994	.995	.992	-

4.7 Residuals of PVC

4.7.1 Frequencies and Spatial Structures After calibrating the vegetation indices it is possible to ascertain the difference between the transformed index and PVC for every 50-m grid square in the area analyzed; it can be interpreted as the residual of PVC. Table 4 displays the frequency distribution of the residuals for the three indices in a non-equidistant form. If the accuracy of PVC is put at ±5 (%) points (cf. section 3), this results in PVC being largely correctly predicted for 48.7% (and thus for almost half) of the 12,585 squares in the settled area of Heidelberg by TNDVI; the standard deviation is 12.0 percentage points. For TPVI and $TPC2_M$ somewhat less favorable values result: standard deviation is 12.7 points for TPVI and 12.2 points for $TPC2_M$.

Table 4: Residuals of the PVC

Residual	TNDVI	TPVI	$TPC2_M$
$-100 \le x < -30$	1.7	1.8	1.5
$-30 \le x < -20$	2.6	2.9	3.1
$-20 \le x < -15$	3.4	3.7	4.1
$-15 \le x < -10$	6.1	6.6	7.3
$-10 \le x < -5$	10.0	11.0	11.5
$-5 \le x \le 5$	48.7	47.1	45.3
$5 < x \le 10$	13.6	12.6	13.9
$10 < x < 15$	7.8	7.0	7.2
$15 < x \le 20$	3.4	3.5	3.5
$20 < x \le 30$	1.9	2.5	1.8
$30 < x \le 100$	0.8	1.2	0.9

Each of the residuals displays a very heterogeneous structure spatially; sharp differences between adjacent grid squares sometimes occur, in sign and/or value. Nevertheless, linear and extensive structures of different kinds and degrees can be recognized in all residuals. Fundamentally, it has to be stated that (vegetation) conditions in the central part of the area analyzed are clearly better reproduced, especially by TNDVI and TPVI, than those of the peripheral areas. This has to be evaluated as a very positive result, especially in terms of the application of vegetation indices in environmental planning, since those regions with a lower percentage of surface vegetation--for which a pressing need for action in ecological terms exist--are more concentrated in the central parts of a city than on its periphery.

For the residuals of individual indices, different causes can be cited:

- Radiometric errors in the geocoded TM data
- Changes in the vegetation cover
- Dependency of the indices on shadow formation
- Dependency of the indices on surface materials

4.7.2 Radiometric Errors Geocoded satellite photography data usually contains, despite systematic restoration, not only geometric errors but radiometric ones as well, which can reach considerable magnitude. Thus, for example, according to FORSTER (1985b, 147), in Landsat MSS data from the urban area of Sydney, only about 50% of the (digital) gray values of a pixel are determined by spectral signals from the corresponding instantaneous field of view (IFOV); the other 50% come from adjoining IFOVs. Even if the findings for these values for Landsat TM data were to be clearly more favorable in optimal atmospheric conditions in the visible and infrared sectors (something which cannot be judged from the EOSAT scene available), this still represents an important source of error in data evaluations in (heterogeneous) urban areas.

In spite of the problems of proof, a clear indication of the existence of radiometric errors can be established in the EOSAT scene. On the basis of 50-m grid squares, a significant correlation of r=0.448 (r^2=0.2) exists between (a) the residuals of PVC that result from an estimation using TNDVI and (b) the deviation of the PVC from the average PVC of its four directly adjoining squares (PVC_4). Figure 3 depicts the isofrequencies of the value pairs for the reduced area of [-40,40] x [-40,40], in which 96.0% of the pixels fall (n=12,193). The isofrequencies given here are somewhat rounded off, since each is the average of 3 x 3 values. They are arranged more or less in the shape of a wide ellipse set diagonally to the coordinate axes with a slope of 1; here, the value pairs in (0,0) display a clear maximum. The value of the residual of PVC tends to be higher, the more a grid square differs from its surroundings in terms of vegetation. The sign of the correlation is not surprising here: the residual tends to be positive (i.e., vegetation is overestimated by TNDVI), if the difference between PVC and PVC_4 is negative (i.e., the surroundings are "greener" than the grid square itself).

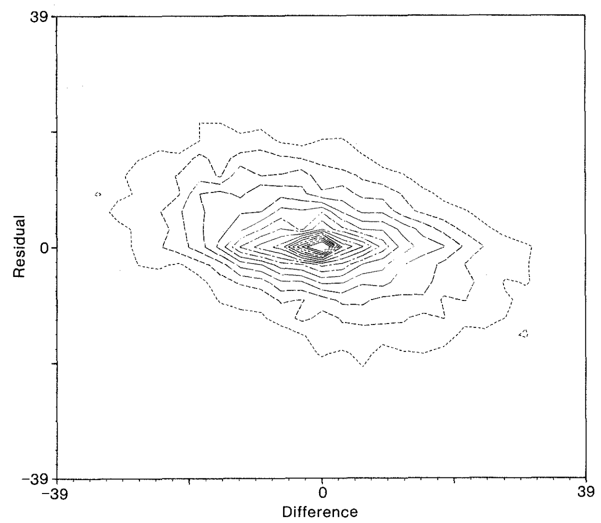


Figure 3: Isofrequencies of the difference ($PVC-PVC_4$) and the residual (TNDVI-PVC), distance of the isofrequencies: 2, maximum in (0,0): 34

This correlation is largely to be attributed to radiometric rather than geometric errors. Even though the geocoded EOSAT scene still contains geometric errors despite careful correction (for one, because scan gaps affect all raw TM scenes), which can also trigger the surroundings effect, these errors are not a sufficient explanation. If the correlation is made between residual (TNDVI-PVC) and difference (PVC_4-PVC) for the ESA scene of 17 August 1987: (a) without eliminating the doubled lines or realigning out-of-place line segments prior to correction, and (b) with the choice, in the correction process, of very unreasonable parameters for the geometric quality of the geocoded subscene (nearest neighbor as resampling procedure, 5 GCPs from a topographical map of 1:50,000), no strong correlation results (r=0.445). For the other two indices (PVI/$PC2_M$), a significant correlation of r^2=0.15/r^2=0.26 between the corresponding residual and the PVC difference also results in the EOSAT scene. Here, however, it still needs to clarified why the percentage of explained variance of the residuals fluctuates to such an extent between the indices.

4.7.3 Changes in the Vegetation Cover Differences in vegetation structures between the aerial photographic referent and those found at the time the satellite pictures were taken are not limited to agriculturally utilized areas. They also extend to areas where the vegetation cover really changed between June 1986 and August 1987. This includes gains in vegetation through new plantings and new seedings and losses in vegetation through building construction or deforestation measures due to a lack of tree vitality. Such areas inevitably in part display strong residuals in all indices, which are not to be judged as errors in the processing of the satellite photographs.

4.7.4 Dependency on Shadow Formation The indices display markedly different values in sloping areas than in level areas when these areas possess quantitatively and qualitatively otherwise largely identical vegetation structures. This effect is the strongest for PVI and $PC2_M$, but can also be demonstrated for NDVI. In general, all indices tend to overestimate vegetation on slopes with eastern and/or southern exposure and underestimate it on slopes with northern and/or western exposure. In the gray-value depiction of NDVI (Fig. 2), this error correspondence is clearly seen in quite a number of areas with valleys, where both slopes are completely covered by vegetation: slopes with southern and/or eastern exposure display higher NDVI values (the degree dependent on steepness) than the slopes with northern and/or western exposure on the other side of the valley. The main cause for the variance of the indices is represented by the angle of sunshine; it was ca. 51° from ESE at the time of filming, which led to considerable shadow formation.

Even though shadow formation only occurs extensively in sloping areas (i.e., in the TM pixel area), it probably sometimes also occurs in level areas--primarily as the result of houses and trees--to such an extent that notable residuals of PVC are caused. The indices in the present study were calibrated using aerial photographic data that, because shaded surfaces are usually correctly classified visually, are hardly affected by this problem. Thus, one can assume that the calibrated indices within the vegetation areas have already taken into account a certain percentage of shadows. For this reason, it is probable that residuals caused by shadows primarily arise where such areas deviate to a considerable extent from the general average.

4.7.5 Dependency on Surface Materials The vegetation indices are--to varying degrees--dependent on the specific composition of the surface materials in the relevant IFOV. This holds not only for areas close to nature or used agriculturally (as has been shown in numerous studies), but also in urban areas. Dependent upon local conditions, identical vegetation conditions (i.e., constant PVC) can be accompanied by sharp differences in the indices. This is largely carried over onto the calibrated indices and leads to considerable error. According to gray-value analysis of selected pixels from the entire subscene of Heidelberg, which can largely be considered pure pixels, all indices display considerable variation around the respective medians x_1 or x_2, both within vegetation (conifers/deciduous trees, bushes, lawns, etc.) and within sealed surfaces (asphalt, concrete, roof tiles, bitumen, asbestos cement, etc.). Within the types of vegetation, PVI displays the largest (relative) variation; within the sealed types of areas, this holds for NDVI. This agrees with the findings of other studies. $PC2_M$ is somewhere inbetween.

4.8 Indices in Test Areas at the Quadrangle Level

A basic factor to be taken into account in the registration of environmental indicators in urban areas is the character of the spatial unit of reference in municipal environmental planning. Important here are not only (square) grid areas, but also administrative units, which generally are polygonal in shape. With digital processing of TM data as the procedure of registration used, given the most favorable of cases, (building) quadrangles represent the lowest administrative unit for which the PVC can be ascertained by means of a vegetation index with sufficient accuracy. To investigate the nature of this accuracy (and for the sake of more extensive studies in the TM subpixel region that have not yet been completed), three test areas were initially selected within Heidelberg (Handschuhsheim, Neuenheim, Town Center) that display different structures of use and vegetation. Here, data was collected, according to the dominance principle in a 2-m grid, on the most important surface materials (trees/bushs, lawns, soil without vegetation, water, asphalt/concrete/cobble- and flagstones, roof tiles, bitumen/asbestos cement). This was once again done through the visual evaluation of color infrared aerial photographs. The test areas encompass 116 hectares (ha); the arithmetical mean of their PVC is 32.0. Each of the grid areas (290,000 in total) was then assigned to exactly one quadrangle, with the middle of the streets forming the boundaries. Quadrangles with an area under 0.33 ha were (a) excluded, if they only were partly located in a test area; or (b) joined to a neighboring test area, if they were located completely within a test area. By way of illustration, Fig. 4 shows the 38 quadrangles taken into account in the Handschuhsheim test area.

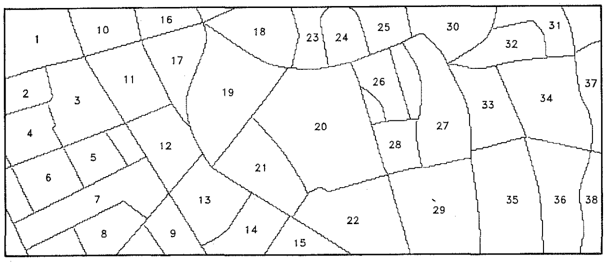


Figure 4: Quadrangle in the test area Handschuhsheim

The following calculation was based on 99 quadrangles, with an average size of 1.1 ha. PVC was determined for every quadrangle in two ways: a reference value was created by aggregating the (aerial) categories tree/bush and lawn; and it was roughly estimated on the basis of the averages (weighted according to area) of the TNDVI data, which display a geometric resolution of 25 m. The correlation of r=0.978 between the two variables is very high. However, it has to be taken into consideration that the test areas already displayed a higher correlation on the 50-m grid level (r=0.939) than the settled area taken as a whole (r=0.923). The standard deviation of the residuals is only seven percentage points in the test areas, and for 77 of the 99 quadrangles, the estimated error is less than five percentage points. This has to be evaluated as a very good result and underscores the potential uses of satellite remote sensing in municipal environmental planning (cf. MEISSNER et al. 1984). Very favorable results are also produced for the other indices (TPVI, $TPC2_M$).

5. SUMMARY

Even though satellite remote sensing represents a favorable procedure for registering relevant, spatially related, environmental data in terms of expenditures of both time and money, it has only rarely been used for planning purposes in urban areas. This is usually justified in terms of the unsatisfactory geometric resolution of present satellite recording systems, which lead to a (too) high degree of error in data evaluation. The goal of the present study has thus been to establish the accuracy with which the percentage of vegetation cover (PVC) in an urban area can be ascertained using vegetation indices on Landsat-TM data. For the settled area of Heidelberg (31.5 km^2), the correlation between each of the calibrated vegetation indices (NDVI, PVI, SAVI, PC2) and the aerial photographic referent made on the basis of 50-m grid squares is on the level of $r=0.92$, which can be held to be a notable result. On the basis of 99 quadrangles in selected test areas of Heidelberg, there is even a correlation in the range of $r=0.97/0.98$ between the indices and PVC. In a spatially differentiated examination, considerable differences emerge among the indices, which we are already acquainted with from other studies: PVI displays comparative advantages in the lower PVC region and NDVI does the same for the upper PVC region. After calibration, SAVI is practically identical to NDVI; thus it is unlikely to present classical indices with much competition. After masking the settled area prior to the principal component transformation, PC2 turns out to possess high quality as a vegetation index, quantitative evidence of which had previously been lacking (from TM data). In spite of this, residuals for all indices display considerable size; standard deviation ranges in all cases in the area of 12/13 percentage points. Even if these values actually turn out to be more favorable, given that the reference data do not always reproduce the vegetation cover at the time of the satellite filming, the goal still must be to improve recording accuracy. This requires, on the one hand, not only the geometric but also the radiometric restoration of TM data; it also requires lessening the dependency of all indices on the composition of surface materials and on shadow formation. Multitemporal and -sensoral data evaluation and the integration of contextual and external information represent means toward this goal; however, extensive research is still needed in these areas. Generally speaking, the results of the study demonstrate that--contrary to a widely held view (cf. BMFT 1990)--even on the municipal level, data from satellite photographs could already be used today for environmental planning.

REFERENCES

ACHEN, M., 1992. Untersuchungen über Nutzungsmöglichkeiten von Satellitenbilddaten für eine ökologisch orientierte Stadtplanung. Heidelberger Geographische Arbeiten, 91.

BMFT, 1990. Empfehlungen der Arbeitsgruppe "Fernerkundungsanwendung für das Forschungsgebiet Ökologie der Landoberflächen" zur Forschungsförderung. Bonn (Interner Bericht).

CURRAN, P., 1980. Multispectral remote sensing of vegetation amount. Progress in physical geography, 4: 315-341.

EHRHARDT, H., 1990. Modellorientierte Entzerrung von Thematic Mapper Rohdaten. DLR-Forschungsbericht 90-55. Köln.

FORSTER, B.C., 1983. Some urban measurements from Landsat data. Photogramm. Eng. remote Sens., 12: 1693-1707.

FORSTER, B.C., 1985a. Principle and rotated component analysis of urban surface reflectances. Photogramm. Eng. remote Sens., 51: 475-477.

FORSTER, B.C., 1985b. An examination of some problems and solutions in monitoring urban areas from satellite platforms. Int. J. remote Sens., 6: 139-151.

HÜBLER, K.-H., 1986. Bodenschutz durch bessere Planungsgrundlagen ? In: Bodenschutz: Räumliche Planung und kommunale Strategien. Bonn, 121-130.

HUETE, A.R. / JACKSON, R.D. / POST, D.F., 1985. Spectral response of a plant canopy with different soil backgrounds. Remote Sens. Environment, 17: 37-53.

HUETE, A.R., 1988. A Soil-Adjusted Vegetation Index (SAVI). Remote Sens. Environment, 25: 295-309.

INGEBRITSEN, S.E. / LYON, R.J.P., 1985. Principal component analysis of multitemporal image pairs. Int. J. remote Sens., 6: 687-696.

KENNEWEG, H. , 1975. Objektive Kennziffern für die Grünplanung in Stadtgebieten aus Infrarot-Farbluftbildern. Landschaft+Stadt, 1: 35-43.

KERL, I., 1989. Flächendeckende Vegetationserfassung der Stadt München mit Hilfe von Satellitendaten. In: SCHLICHER, M. / FRITSCH, D. (Ed.). Geo-Informationssysteme, Karlsruhe, 219-227.

MEISSNER, B. / DOBBRICK, K. / MURNIER, C., 1984. Satelliten-Fernerkundungs-Daten für die städtische Grünübersicht ? In: Berliner Geow. Abh., C3, 147-151.

OLBRICH, G., 1990. Abgrenzung des Siedlungsbereichs von Heidelberg mit Hilfe von Landsat 5 Thematic Mapper Daten. In: DLR-Mitteilungen 91-09, 50-52.

PERRY, C.R. / LAUTENSCHLAGER, L.F., 1984. Functional equivalence of spectral vegetation indices. Remote Sens. of Environment, 14: 169-182.

RADERMACHER, W., 1989. Statistisches Informationssystem zur Bodennutzung. In: MARKWITZ, W. / WINTER, R. (Ed.). Fernerkundung. Karlsruhe, 19-30

RICHARDSON, A.J. / WIEGAND, C.L., 1977. Distinguishing vegetation from soil background information. Photogramm. Eng. remote Sens., 43: 1541-1552.

SOLIDAY, B.L., 1986. Identification of data redundancies in vegetation index models: An application of Landsat spectral reflectance digital data. IGARSS, Vancouver, 3: 1339-1342.

USBECK, B., 1989. Klassifizierung städtischer Nutzflächenarten aus multispektralen Luftbildern. In: Hallesches Jahrbuch für Geowissenschaften, 14, 65-73.

WARDLEY, N.W., 1984. Vegetation index variability as a function of viewing geometry. Int. J. remote Sens., 5: 861-870.