

Article

Spatio-Temporal Assessment of Tuz Gölü, Turkey as a Potential Radiometric Vicarious Calibration Site

Vincent O. Odongo ^{1,2,*}, Nicholas A. S. Hamm ² and Edward J. Milton ¹

¹ Geography & Environment, University of Southampton, Southampton, Hampshire SO17 1BJ, UK; E-Mail: e.j.milton@soton.ac.uk

² Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, P.O. Box 217, 7514 AE Enschede, The Netherlands; E-Mail: nick@hamm.org

* Author to whom correspondence should be addressed; E-Mail: v.o.odongo@utwente.nl; Tel.: +31-5-348-74524; Fax: +31-5-348-74554.

Received: 15 November 2013; in revised form: 5 March 2014 / Accepted: 10 March 2014 /

Published: 20 March 2014

Abstract: The paper provides an assessment of Tuz Gölü, a site in Turkey proposed for the radiometric vicarious calibration of satellite sensors, in terms of its spatial homogeneity as expressed in visible and near-infrared (VNIR) wavelengths over a 25-year period (1984–2009). By combining the coefficient of variation (CV) and Getis statistic (Gi*), a spatially homogenous and temporally stable area at least 720 m × 330 m in size was identified. Analysis of mid-summer Landsat Thematic Mapper (TM) images acquired over the period 1984–2009 showed that the hemispherical-directional reflectance factor of this area had a spatial variability, as defined by the CV, in the range of 0.99% to 3.99% in Landsat TM bands 2–4. This is comparable with the reported variability of other test sites around the world, but this is the first time an area has been shown to have this degree of homogeneity over such a long period of time.

Keywords: vicarious calibration; validation; spatial variation; radiometry; Getis statistic

1. Introduction

Accurate satellite sensor calibration is vital if data products derived from Earth observation (EO) are to be reliable sources of data for decision support systems in environmental management. This is especially important in studies involving monitoring change over time or combining data from

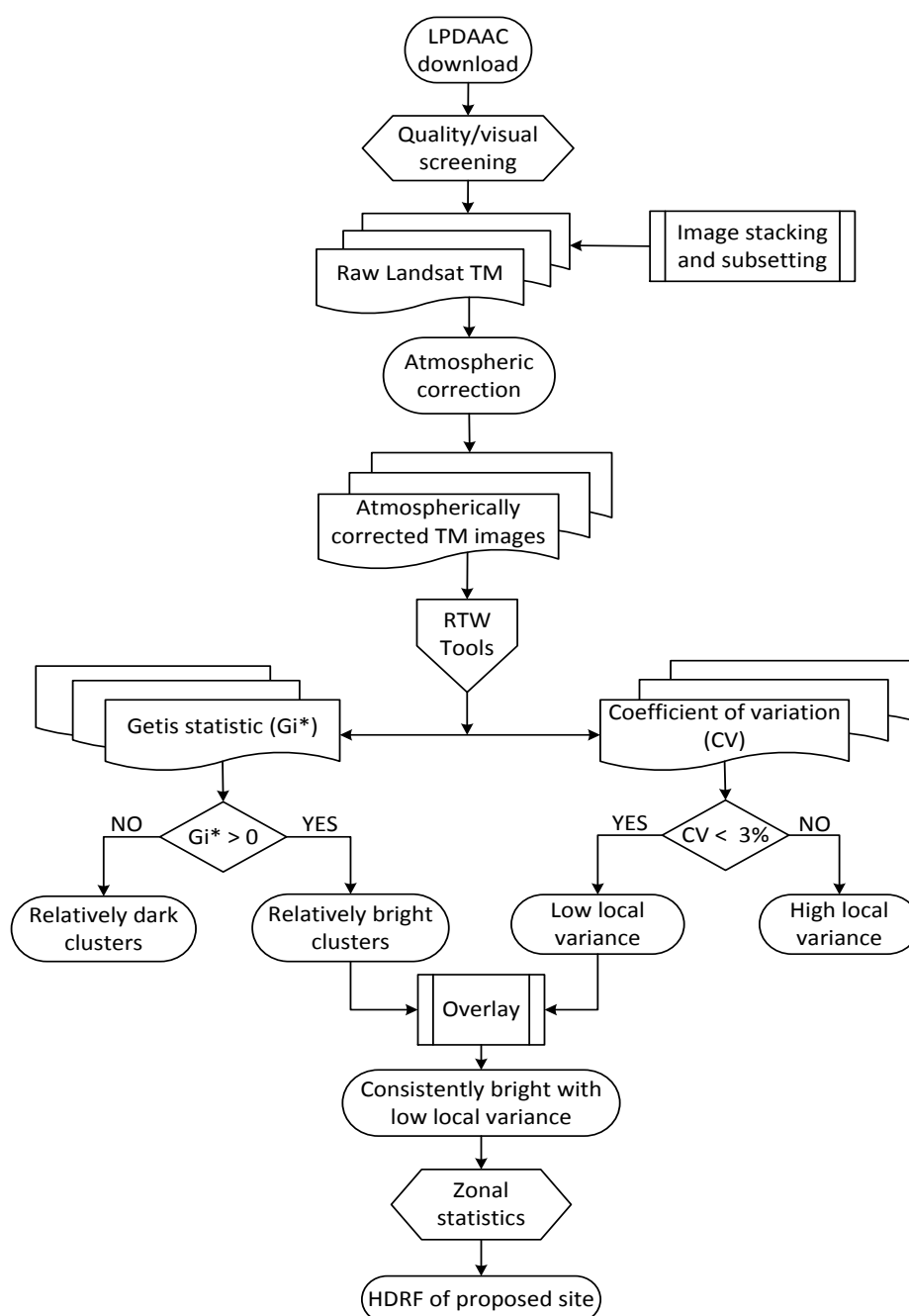
different sensors or platforms to provide accurate and reproducible geophysical and biophysical variables [1]. Radiometric sensor calibration, maintaining that calibration over the lifetime of the system, and the provision of traceability to international standards are all vital steps to ensure confidence in EO data products [2,3]. Vicarious calibration (VC) using ground targets has become established as an important method of post-launch radiometric calibration, and the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV) has endorsed eight terrestrial sites for the post-launch calibration of space-based optical imaging sensors, including Tuz Gölü, a seasonal salt lake in Turkey (latitude 38.50 °, longitude 33.20 °) [4,5]. This paper presents an independent evaluation of Tuz Gölü as a VC site, with an emphasis on its long-term spatial homogeneity. Desirable characteristics of VC sites are well-established in the literature, and include high spectral reflectance (above 30% in all bands), spatial homogeneity, an extensive flat, smooth surface and little or no vegetation [6–8]. Sites with all or most of these properties that are also at high altitude and in arid climatic zones are favoured most as the effect of the atmosphere on the ground-leaving radiance is minimized. High altitude salt pans, dry lake beds and sandy desert areas are most commonly chosen as VC sites [6–8].

The spatial homogeneity of VC sites is crucial to their effectiveness and their suitability for use with satellite sensors with different ground resolution cell sizes. Visually homogeneous surfaces such as dry lake beds or salt pans have complex geomorphology in which flows of water and sediment interact in a more or less organized fashion to produce highly dynamic, spatially-differentiated environments. The terrain units which make up these environments are themselves dynamic, so the spatial homogeneity of a defined area will vary in different wavelengths and over time. Selection of an appropriate VC site thus involves searching for homogeneity in all three domains of interest: spatial, spectral and temporal. Having selected a candidate site based on these criteria it can then be investigated on the ground to test whether it is also stable in the angular domain (Bidirectional reflectance distribution function, BRDF) [9,10] and whether local infrastructure makes it suitable as a long-term instrumented site.

The most common statistic used to assess the spatial homogeneity of a VC site is the coefficient of variation (CV) [11]. The CV (see Section 3.3) gives a measure of the variability of the reflectance for a given waveband within a defined area of an image. This area may be a local window which is moved across the image. Alternatively the area may be a subset of the target or the entire target. The CV tends to change with increasing pixel size (coarser resolution), although the exact pattern varies between scenes [12]. Furthermore, for a given pixel size, the CV may vary according to the window or subset size, neither of which are standardized in the literature. For example, the authors of [13] refer to an area of 20 pixels and the authors of [14] refer to a 3 km × 3 km area for Satellite Pour l'Observation de la Terre High Resolution Visible (SPOT HRV) and Landsat Thematic Mapper (TM). In their study of the Lunar Lake Playa site, the authors of [1] investigated CV over window sizes varying from 3 × 3 to 9 × 9, 30-m pixels and settled on a 5 × 5 window as a compromise between excessive detail and over-generalization. Based on previous studies, the authors of [1] considered a site to be homogeneous if the CV within the chosen window size was 3% or less, and that criterion will be adopted in this paper. Higher CVs have sometimes been considered acceptable, especially when the CV is computed within a larger subset. The temporal stability of the VC site can be assessed by determining the extent to which this spatial-CV varies over time [13]. If it remains within the 3% threshold, the site is

considered temporally stable. An alternative approach was adopted by [11], who actually calculated CV for a given pixel over time.

Figure 1. Schematic approach used to identify a spatially uniform, temporally invariant site for sampling reflectance for vicarious calibration. Landsat TM images were acquired from NASA Land Processes Distributed Active Archive Centre (LPDAAC). RTW (Robin Wilson Tools) was used to compute Getis statistic (G_i^*) and Coefficient of variation (CV).



The CV is important since it allows assessment of whether an area meets a criterion for overall variability. Nevertheless, it is of limited value for assessing the spatial homogeneity of a target [1,15,16]. Two or more surfaces may have different means and standard deviations yet have the same CV. Hence adjacent areas with similar CVs may actually have quite different properties. Furthermore, the CV is

not a measure of spatial association. This has led some authors (e.g., [1,15,16]) to propose the use of the Getis, G_i^* , statistic [17–19]. The G_i^* statistic is a local indicator of spatial association (LISA) that gives a measure of clustering. When similar values are clustered, this is indicative of spatial dependence and considered indicative of spatial homogeneity [17–19]. This is useful, since the G_i^* statistic can be used to identify clusters of similar reflectance values. The term “uniformity” is sometimes used in the VC literature. For this paper the term “homogeneity” is used in preference to “uniformity” to describe this characteristic since its meaning is more clearly defined statistically in terms of stationarity and spatial dependence [17].

The objective of this research was to identify high reflectance, low variability, spatially homogeneous areas within Tuz Gölü, and to study their persistence through time using a 25-year time series of Landsat TM data. The authors of [4] observed low reflectance (<20%) in the short wave infra-red (SWIR) and recommended the site suitability be restricted to the VNIR. Hence, in this study focus is on the visible and near-infrared (VNIR) bands (Landsat TM Band 2–4), which are the most-used bands for land applications. The results of this analysis can be used to target specific areas for field investigation. Figure 1 is a flowchart of the procedure, which is described in detail below. The research presented in this paper extends the previous studies at the same site by [16] in three ways. First, it is based on fine spatial resolution data from Landsat TM, and therefore allows a more detailed assessment of the site and its suitability for vicarious calibration of a range of sensors. The results are directly relevant to sensors with comparable spatial resolution to Landsat TM (e.g., SPOT, Disaster Monitoring Constellation (DMC)), but they also provide an insight into the spatial heterogeneity of the site at much larger pixel sizes and therefore have relevance to sensors with coarser resolution (e.g., Medium Resolution Imaging Spectrometer (MERIS) and Moderate Resolution Imaging Spectro-radiometer (MODIS)). Second, the study is based on a much longer time series than earlier work, and therefore provides evidence of the long-term suitability of the site. The opportunity to use the same VC site throughout the lifetime of a satellite mission has important practical and scientific advantages. Finally, the paper approaches the problem of VC site characterization from a broader perspective than earlier studies (e.g., [1,15]), as it emphasizes local variability and clustering as related aspects of a general approach to VC site characterization.

2. Site Description

Tuz Gölü is an ephemeral lake in the central Anatolian region of Turkey, approximately 150 km south-east of the city of Ankara (lat. 38.50 °, long. 33.20 °) at an altitude of 905 m above mean sea level (Figure 2). For most of the year Tuz Gölü is a shallow, saline lake approximately 80 km long and 50 km wide, but in the summer months (July–August) when the temperatures are high and mostly zero precipitation (Figure 3) experienced [20], the lake dries out exposing a salt flat consisting mainly of halite and gypsum evaporites with minor amounts of polyhalite, coelestine and the economically important mineral huntite [21,22]. Beneath the salt crust (1–30 cm thick) there is a layer of unconsolidated muddy sediment up to 25 cm thick comprising gypsum, magnesite, thenardite, polyhalite, aragonite and montmorillonite [22]. Two saltworks operate on the shores of the lake, and the surrounding area is mostly agricultural fields and steppe vegetation [16].

Figure 2. Location map of Tuz Gölü, Turkey.

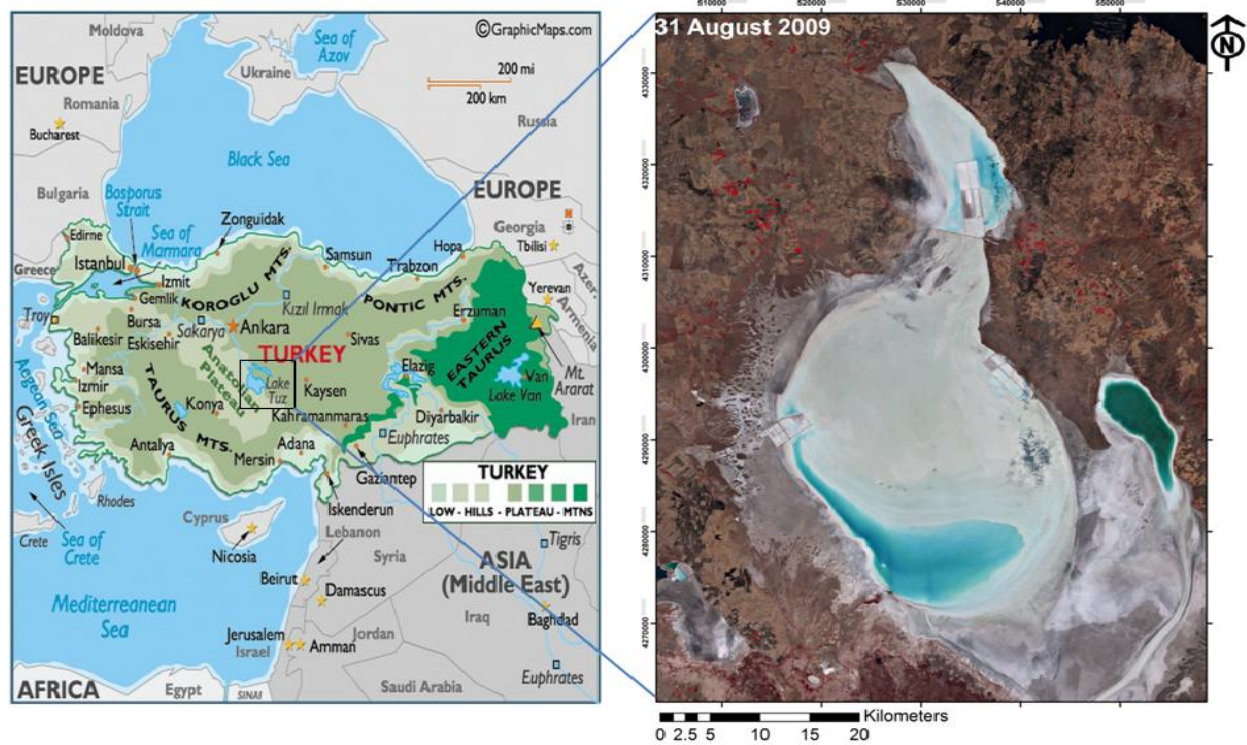
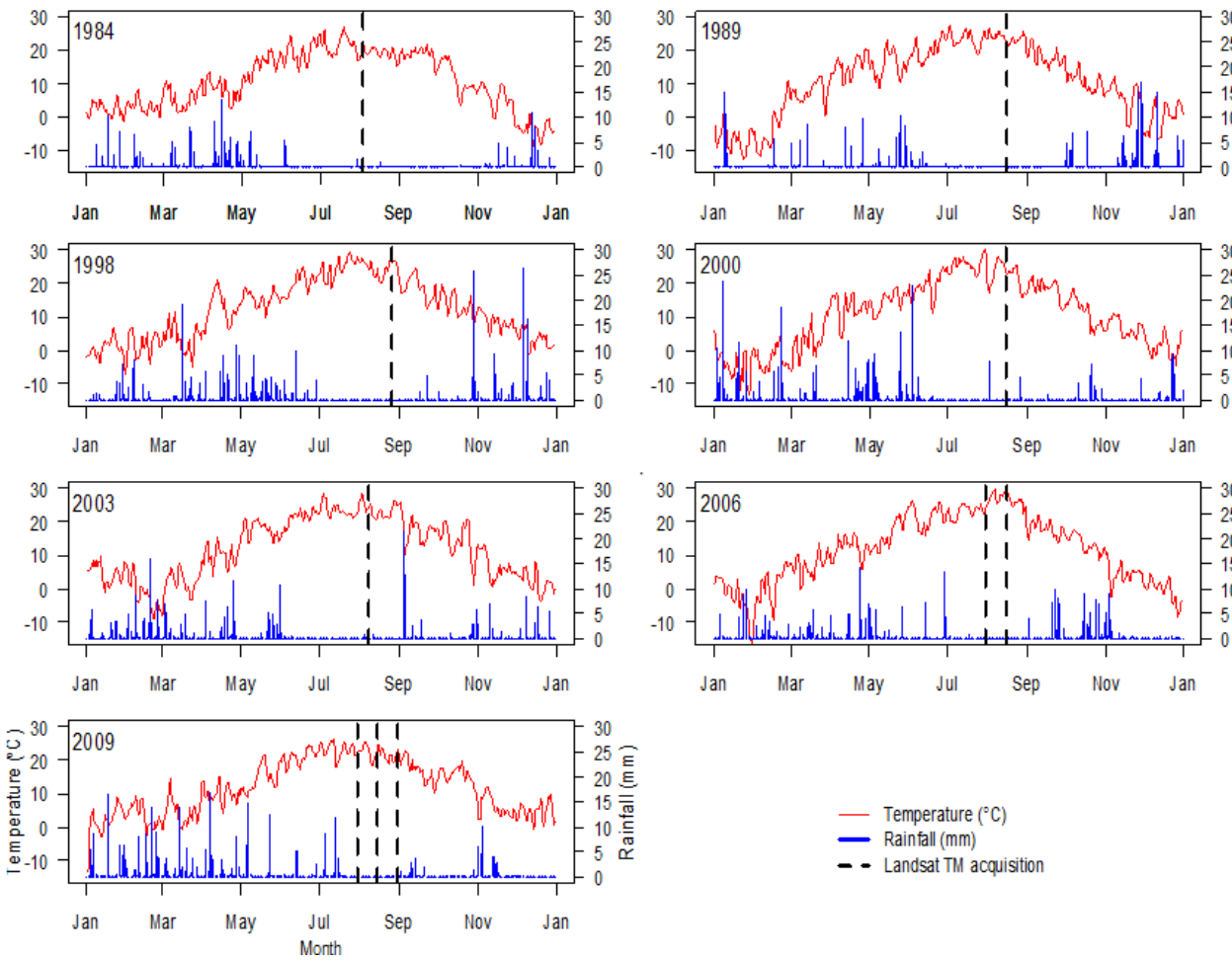


Figure 3. Long-term daily average rainfall and temperature at Tuz Golu.



3. Method

3.1. Data Acquisition and Processing

When assessing the suitability of a VC site, most information is gained from a long time series of data acquired with a fine resolution sensing system. The Landsat TM is ideal for this purpose, having been in orbit since 1984 and providing well-calibrated spectral data in six reflective bands with approximately 30 m spatial resolution [23]. Investigations on the calibration accuracy of the Landsat-5 TM by [23] explain that the accuracy was in the order of 5% during pre-launch and post-launch, assessed through on-board and vicarious calibration. From 1999 onwards, the sensor has been shown to maintain an accuracy within 2% [23]. TM data are freely available and were therefore ideal for assessing the suitability of a candidate VC site.

A time series of cloud-free Landsat TM images for July and August were obtained from the NASA Land Processes Distributed Active Archive Centre (LPDAAC) (Table 1 and Figure 4). The solar zenith angle (SZA) of the images varied between 29.04 and 36.77 degrees, which is unlikely to significantly affect the reflectance of a flat, bare surface. The sensor view angle for Landsat TM is within a few degrees of nadir. The images were radiometrically and atmospherically corrected using the Atmospheric/Topographic Correction for Satellite Imagery (ATCOR-2) program [24]. No concurrent atmospheric data were available from the site so the MODTRAN mid-latitude summer atmospheric profile was used, together with the aerosol type “desert” for all images. Of the standard aerosol types offered by radiative transfer models, “desert” is most likely to represent the conditions found at Tuz Golu. Experiments with ATCOR-2 showed that changing the aerosol type (e.g., to continental) had no significant effect upon the results.

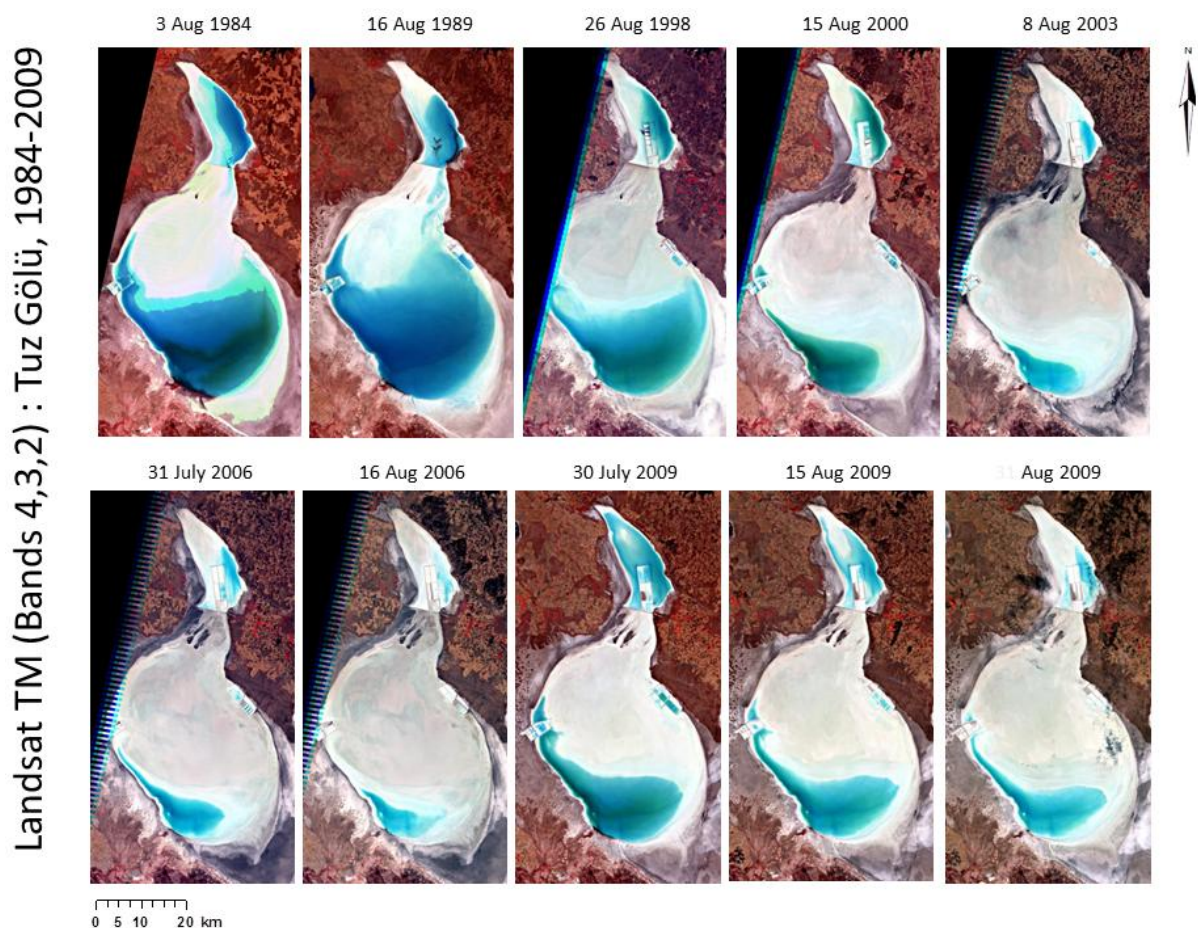
Table 1. Landsat TM data used in the study obtained from NASA Land Processes Distributed Active Archive Centre (LPDAAC).

Date (YYYY-MM-DD)	Time (GMT)	Cloud Cover (%)	Solar Zenith Angle (degrees)
1984-08-03	7:50:51	0	33.39
1989-08-16	8:01:06	0	35.28
1998-08-26	8:00:04	0	36.77
2000-08-15	7:59:06	0	34.51
2003-08-08	7:58:19	0	33.03
2006-07-31	8:14:43	0	29.04
2006-08-16	8:14:54	0	32.31
2009-07-30	8:16:43	0	29.55
2009-08-15	8:16:57	0	32.73
2009-08-31	8:17:12	4	36.38

The resulting hemispherical-directional reflectance factor (HDRF) images were masked to eliminate the area outside the lake area. Landsat TM Bands 1–4 were the primary focus of this work because they cover the bands used in most land applications (e.g., the bands used in NDVI, *etc.*). The other short-wave infra-red (SWIR) bands of Landsat TM (Bands 5 and 7) are not needed for many vegetation-related applications, which predominantly use NDVI and similar indices. Minerals found

on salt pans such as Tuz Golu have absorption features in the SWIR, so such sites are most suited to VC of VNIR wavelengths. The original DN values in Landsat TM band 1 were saturated over the highly reflective salt pan during the period (July and August) when Tuz Golu would be used for VC so this band was not used in the analysis. Analysis was restricted to the other VNIR bands 2–4. Wilson [25] provided a suite of IDL/ENVITM programs to compute the Getis statistic and coefficient of variation over variable size windows within a masked image. Most of the other spatial analysis was performed within ESRI ArcGISTM version 9.3.

Figure 4. Landsat TM colour composite images (Red, Green, Blue = TM 4, 3, 2) showing Tuz Gölü for the months of July and August from 1984 to 2009. The blue patches represent shallow water.



3.2. Analysis of Clustering (G_i^*)

The variogram, Moran's I or Geary's c are global statistics commonly used to provide a measure of the spatial structure or spatial autocorrelation over an area. In contrast, the so called local indicators of spatial association (LISA) provide a measure of spatial association within a restricted window or sub-area. The Getis statistic, G_i^* , is part of this family and gives a measure of clustering [17–19]. It is defined as (Equation (1)):

$$G_i^*(d) = \frac{\sum w_{ij}(d)x_j - W_i^* \bar{x}}{s[W_i^*(n - W_i^*)/(n - 1)]^{1/2}} \quad (1)$$

where n is the total number of pixels, $w_{ij}(d)$ is a matrix of spectral weights with binary and symmetric weight equal to unity ($w_{ij} = 1$) for all pixels found within distance d of pixel i considered and a weight equal to zero ($w_{ij} = 0$) for all pixels found outside d . $\sum w_{ij}(d)x_j$ is the sum of varying values X (i.e., reflectance of the imagery) within distance d of pixel i (i included), W_i^* is the number of pixels within the distance d (i included). For regular grids, including images, $\sum w_{ij}(d)x_j$ can be defined within a square window (e.g., $d = 1$ is 3×3 window). The aim of this study was to identify suitable sub-areas within Tuz Gölü, hence the surrounding areas were masked to ensure that they did not affect the estimate of the mean (\bar{x}) and standard deviation (s) and, hence the calculation of G_i^* .

Spatial homogeneity refers to the extent to which a surface comprises the same kind of elements or is similar in nature (i.e., exhibiting similar spatial structure in reflectance) in space [6]. Since the G_i^* gives a measure of clustering, it can be used to identify pixels that are bright relative to the mean ($G_i^* > 0$) or dark relative to the mean ($G_i^* < 0$). This was demonstrated by [1] who used the G_i^* to identify clusters of relatively bright pixels on the Lunar Lake Playa, Nevada, USA (LLPN). The broader applicability of the statistic in remote sensing was reviewed by [26] and demonstrated by [27]. The research presented in this paper follows the guidance of [1] and [16] who used the $G_i^* > 0$ to identify clusters of relatively bright pixels.

The G_i^* was determined at a scale of $d = 1$ (3×3 window), yielding a statistic which is a highly localized. This choice of a 3×3 window is in line with several other authors [6,11,16].

G_i^* was computed for each image for TM bands 2–4. G_i^* values greater than zero were then extracted by overlaying all the G_i^* output maps for bands 2–4 for all the years represented in Table 1. The output was a single map showing the areas within Tuz Gölü that exhibit consistently bright clusters in all bands over time.

3.3. Analysis of Local Variability (CV)

The CV gives a standardized measure of variability. It is defined in Equation (2).

$$CV = \frac{s}{\bar{x}} \quad (2)$$

where s is the standard deviation of the spectral measurements and \bar{x} is the mean of the spectral measurements within a pre-defined window (e.g., 3×3). As with G_i^* , the CV was computed in a 3×3 window for all the images shown in Table 1. As for G_i^* , the CV was also calculated for a 9×9 window for comparative purposes. Finally, the CV was calculated for the eventually-identified calibration target.

The CV can, therefore be considered a measure of local variability, as described by [12]. The CV has been used to estimate spatial homogeneity of other VC sites, including Railroad Valley Playa, US [1,6,14], Saharan and Arabian Desert sites [28], La Crau, France [29] and Western Queensland, Australia [11]. With the exception of [11], those studies considered an area to be temporally stable if the CV remained within a threshold (e.g., 3%) over time. For the current research, the surface was considered temporally stable if it had a CV of less than 3% in all bands for all images over the relevant time period. This follows the advice of [7]. As with G_i^* , CV was calculated within a 3×3 window for each of Landsat bands 2–4 and for all the images shown in Table 1.

3.4. Identifying Spatially Homogeneous Areas

The CV is a standardized statistic so adjacent areas may have similar CVs, but different means and standard deviations. The combination with G_i^* was important in this respect since relatively bright clusters of low local variability were identified. Suitable areas for use as a ground calibration target were, therefore, identified according to the following criteria

Criterion 1: $G_i^* > 0$;

Criterion 2: $CV < 3\%$;

Criterion 3: $HDRF > 30\%$;

Criterion 4: Persistence of criteria 1–3 over time.

Together, (1) and (2) identify spatially homogenous areas, whilst (3) satisfies the criteria for bright targets. Where these criteria persisted over time (4), the area was judged to be temporally stable. This was proposed as the most suitable area for vicarious calibration and atmospheric correction activities.

4. Results

4.1. Spatial Clustering

Figure 5 shows examples of G_i^* maps for the 3×3 pixel window for 1984, 2000 and 2009. It is evident that there are bright-clustered areas concentrated at the western middle part, northern limb and at the southern fringes of the lake. Figure 6 shows the bright clusters (3×3 window, $G_i^* > 0$) overlain for all bands and all images given in Table 1 (The complete series of G_i^* maps are provided in Supplementary Figure S1). These constitute 27.4 km^2 which is approximately 2.7% of the whole area. It is also evident that much of the area within Tuz Gölü (approximately 97%) was, according to this criterion, unsuitable. It is clear that the inclusion of images from several dates restricted the area identified as suitable.

Bright clusters in the southern fringes and northern limb of the lake were of less extent compared to the upper middle part. The results showed that there were bright areas that were consistently clustered year-on-year. These areas were characteristically clustered in zonations typical of salt pans and playas [30].

4.2. Local Variability

Figure 7 shows examples of CV maps of the study area for the years 1984, 2000 and 2009. The complete series of CV maps are provided in Supplementary Figure S2. A large area falling in the middle of the lake was found to satisfy the criteria $CV < 3\%$ when all the bands were evaluated for all the years (Figure 8). The total area satisfying this condition was approximately 300 km^2 which is the equivalent of 29.3% of the whole area. The CV time series showed that a large area of the lake has low local variability, with the notable exception of the area in the south-west which in some years had residual shallow waters which caused band 4 to be more variable (Figure 7). Bands 2 and 3 were unaffected, showing that the water was very shallow.

Figure 5. Getis statistic map of Tuz Gölü in TM bands 2–4 for the months of August 1984 (Top), August 2000 (Middle) and August 2009 (Bottom). Black indicates $G_i^* > 0$ and white indicates $G_i^* < 0$.

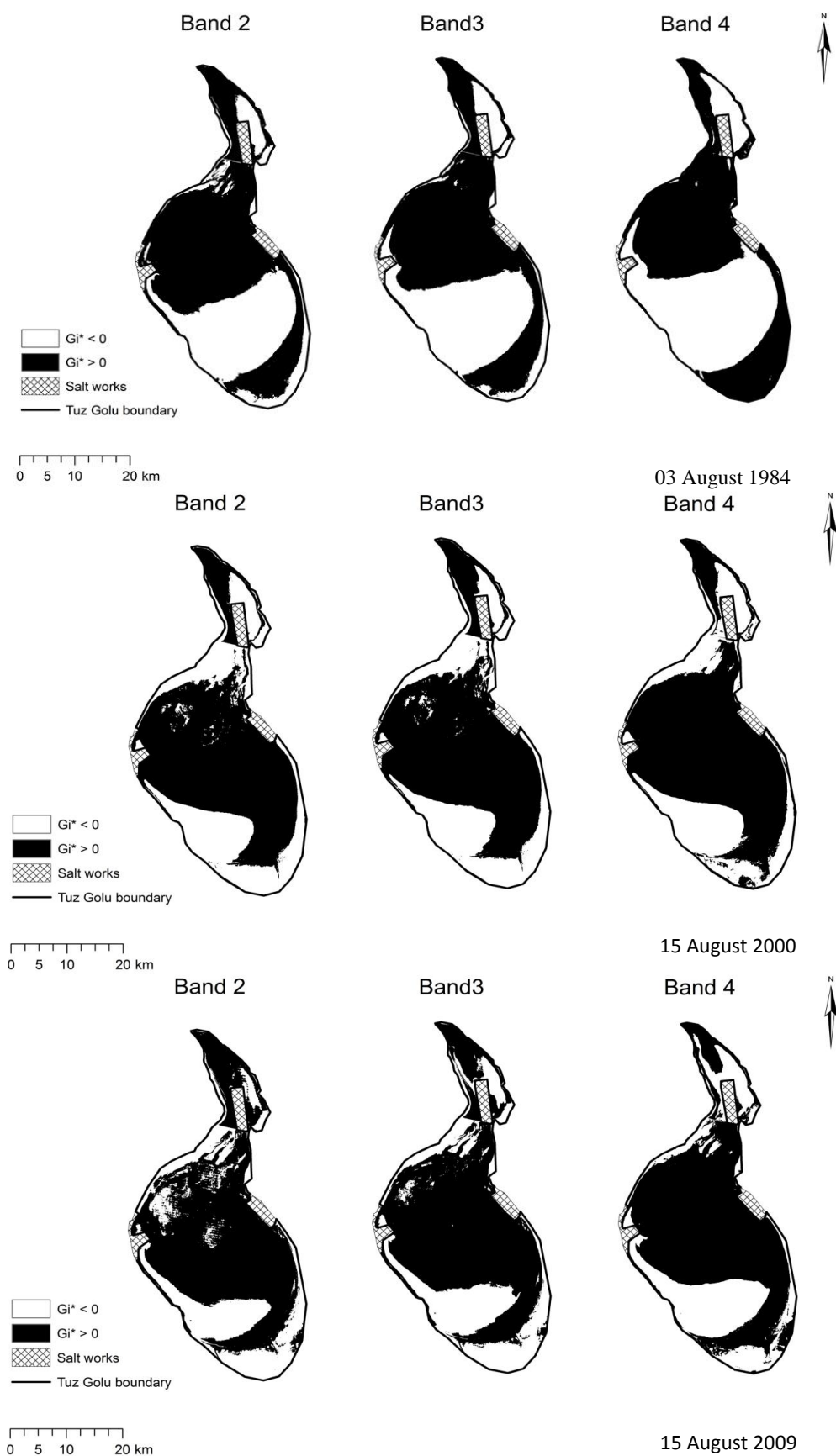


Figure 6. Persistently high reflectance clusters derived from the complete time series of Getis statistic maps. Black indicates $Gi^* > 0$ in all years and white indicates $Gi^* < 0$ in at least 1 year.



Figure 7. CV map of Tuz Gölü in TM bands 2–4 for the months of August 1984 (Top), August 2000 (Middle) and August 2009 (Bottom). Colour ramp goes from black (CV < 3%) to white (CV > 20%).

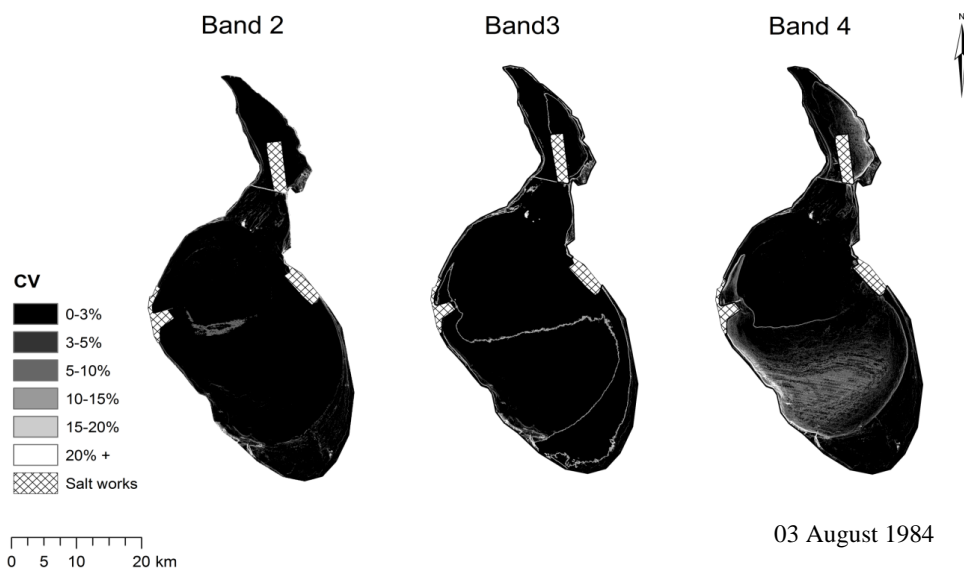
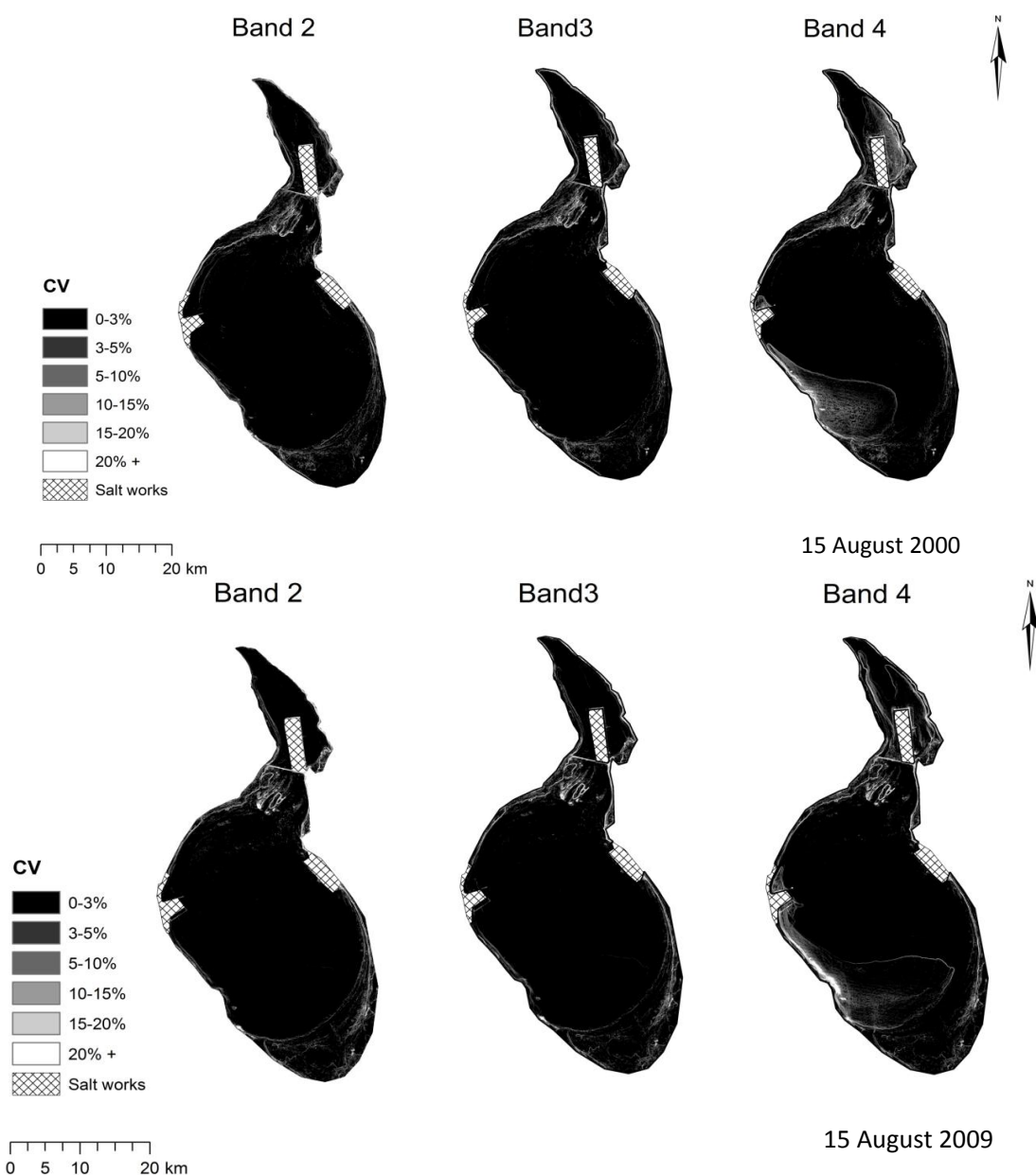


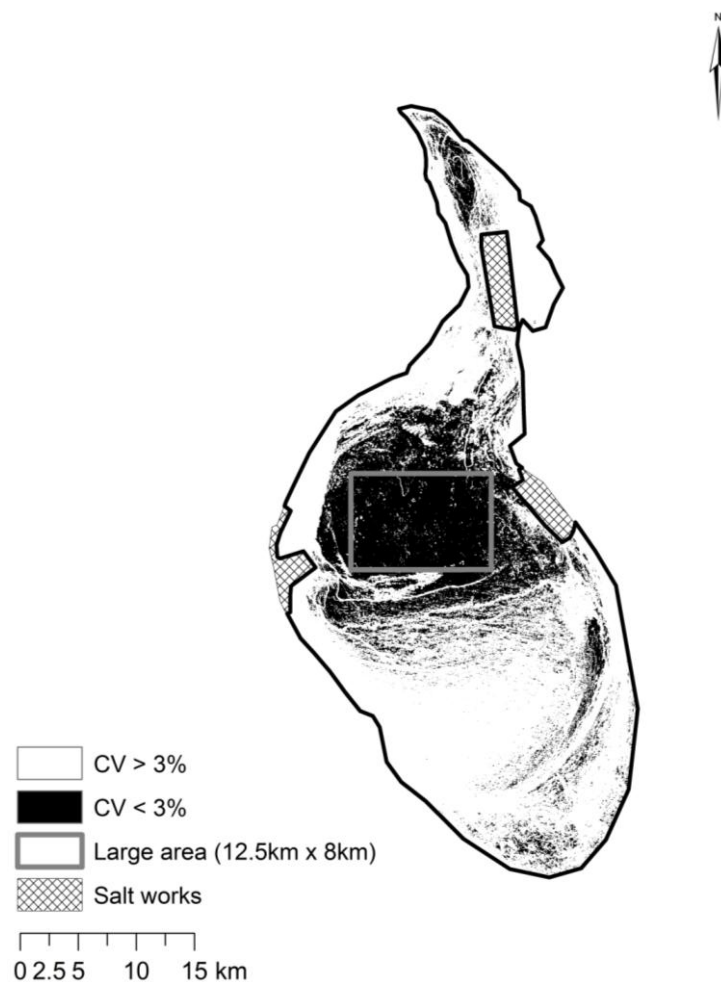
Figure 7. Cont.



4.3. Spatial Homogeneity

Finally, the areas that were persistently bright (Figure 6) and of low local variability (Figure 8) were overlain to identify a spatially homogeneous and temporally stable area. This is illustrated in Figure 9. The visually-identified site satisfying these two criteria is 720 m \times 330 m (proposed site 1), almost 1.5 times larger than the area of La Crau test site in France and Railroad Valley Playa (RRVP) in the USA. When some isolated pixels with CV $>$ 3% were ignored, the area was expanded to 690 m \times 840 m (see Figure 9: proposed site 2). Such isolated pixels tend to be outliers rather than areas with systematically high variability. Ignoring these isolated pixels would be justifiable if it did not lead to a substantial increase in the overall CV for the selected area.

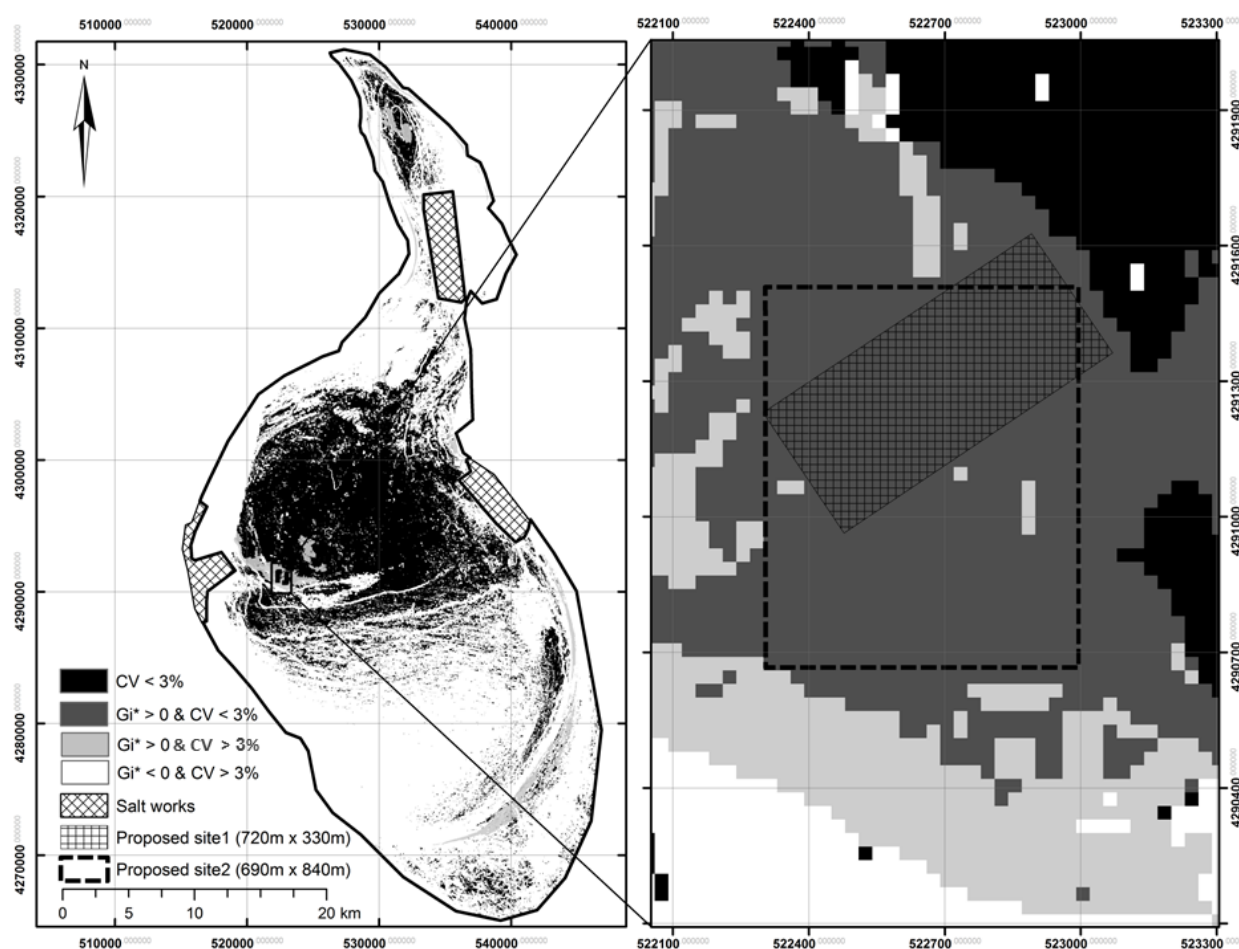
Figure 8. CV map for all bands in all the years. Black indicates $CV < 3\%$ in all years and white indicates $CV > 3\%$ in at least 1 year.



4.4. HDRF of Proposed Site

The mean HDRF of the proposed site was extracted from the radiometrically and atmospherically corrected images. This was accomplished by computation of zonal statistics using the ESRI Spatial AnalystTM tool. This was performed for the $720\text{ m} \times 330\text{ m}$ area. The results are summarized in Table 2 and illustrated in Figure 10, which shows the mean HDRF on each date and its standard deviation. The chosen site is highly reflective in all three TM bands, although it would be unwise to infer too much from the change in HDRF over time as the atmospheric correction applied to the images was based on a standard atmosphere. Hence the absolute reflectance values and therefore the trend through time may be inaccurate.

The spatial variability in the HDRF of the proposed area was not affected by the accuracy of the atmospheric correction. The atmospheric correction included a scattering correction for radiation reflected from neighbouring pixels, but the effect of this on the image statistics was negligible. The CV was computed using Equation (2) for the identified areas for all the years under study. The CV was below 3% for all bands for all years, with the exception of band 4 for the 1989 image, which had a CV of 3.99%. August 1989 was a wet by comparison to other years (see the time series of imagery in Figure 4, which probably led to the relatively high CV in band 4).

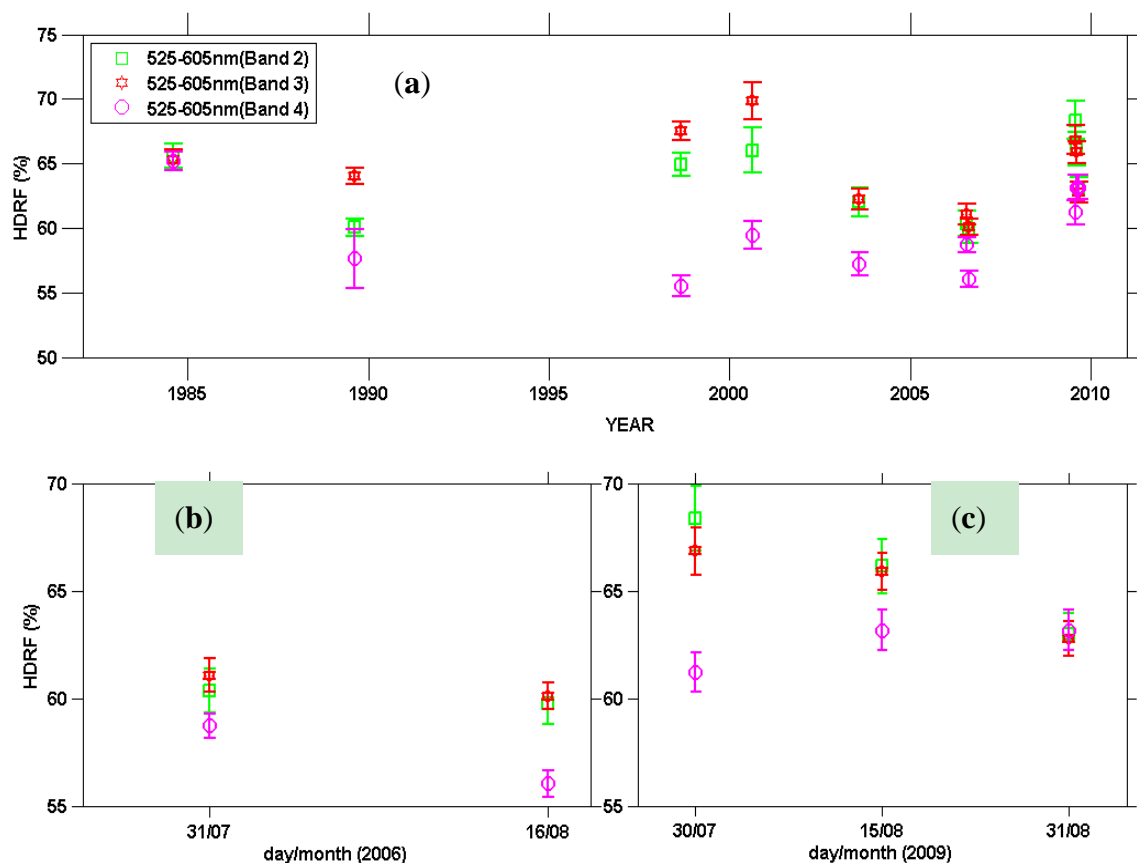
Figure 9. Spatio-temporal stability map for Tuz Gölü showing the proposed calibration site.

As stated in Section 4.3, when a few isolated pixels with a $CV > 3\%$ were ignored a much larger area was identified ($690\text{ m} \times 840\text{ m}$ see Figure 9, proposed site 2). This larger area had a $CV < 3\%$ for all bands for all years (maximum 2.34%), with the exception of band 4 for the 1989 image, which had a CV of 4.96%. The statistics are provided in Supplementary Table S1.

Table 2. Mean, standard deviation (SD) and coefficient of variation (CV) (%) of the HDRF across the three bands within the proposed site 1 (Figure 9) for the years under study.

Date	TM Band 2			TM Band 3			TM Band 4		
	Mean	SD	CV (%)	Mean	SD	CV (%)	Mean	SD	CV (%)
1984-08-03	65.59	0.94	1.44	65.28	0.81	1.24	65.19	0.73	1.12
1989-08-16	60.05	0.66	1.09	64.02	0.63	0.99	57.64	2.30	3.99
1998-08-26	64.92	0.89	1.37	67.54	0.72	1.07	55.52	0.82	1.48
2000-08-15	66.03	1.72	2.61	69.82	1.44	2.07	59.45	1.06	1.79
2003-08-08	62.01	1.08	1.74	62.24	0.80	1.28	57.20	0.90	1.58
2006-07-31	60.34	1.01	1.67	61.06	0.77	1.26	58.71	0.59	1.01
2006-08-16	59.77	0.98	1.64	60.09	0.61	1.01	56.03	0.63	1.13
2009-07-30	68.35	1.49	2.18	66.83	1.12	1.67	61.21	0.92	1.50
2009-08-15	66.13	1.26	1.90	65.88	0.85	1.29	63.17	0.96	1.52
2009-08-31	62.96	1.01	1.61	62.78	0.81	1.29	63.17	0.96	1.52

Figure 10. Satellite derived HDRF plots of Bands 2–4. Error bars show percent standard deviation in HDRF for respective years. (a) Entire series from 1984 to 2009; (b) 2006 HDRF; (c) 2009 HDRF.



For comparison a large contiguous area of pixels was identified for which the CV, as calculated within a 3×3 window, was less than 3% (*i.e.*, most pixels in this area had a $CV < 3\%$ within the coincident 3×3 window). The CV for the whole of this large area was then calculated. This mimics the conventional approach of identifying a suitable site, based on the CV within a local window. The identified area is shown in Figure 8 and the statistics provided in Supplementary Table S2. The CV of this contiguous area was above 3% for most bands across all images. This demonstrates clearly the utility of the Getis + CV method, as compared to the conventional approach of using only the CV, for the identification of a low variability homogeneous area.

5. Discussion

Spatial clustering and local variability of a test site can be affected by different factors such as surface moisture and variation in topography, which in turn create shade effects, and vegetation causing spectral changes [6]. For the case of Tuz Gölü, it might be expected that moisture conditions in the lake would vary year by year thus affecting spatial homogeneity and temporal stability. The results showed that there were bright areas that were consistently clustered year-on-year. These clusters were characteristically patched in zonations typical of salt pans and playas [30]. Possibly, these zonations could be due to sub-environments identifiable by aggregation of salt crusts usually formed in the centre of playas [31].

Many studies involving the characterization of the spatial homogeneity of vicarious calibration surfaces have used only CV (e.g., [6,14,28]). This only provides information on the level of variability within the sub-area or window (Table 3). It does not provide information about whether similar values are clustered. Indeed it may be misleading because two connected surfaces may have equal CV but have different means and standard deviations. It is argued here that the combination of the Gi^* and CV addresses this problem. The CV allows selection of areas with low local variability, whilst the Gi^* allows the selection of relatively bright clusters. When the two statistics are combined they yield a measure of homogeneity. The criteria for homogeneity is specified as $CV < 3\%$ and $Gi^* > 0$. This is important for VC campaigns since it is necessary that homogeneity be upheld across all bands of the sensor and from year to year.

Table 3. Examples of reported temporal CV for VC sites.

Test Site	CV	Sensor	Authors
La Crau, France	5 to 10%	SPOT	[29]
La Crau, France	10 to 15%	ATSR-2	[13]
Railroad Valley Playa, US	1.2 to 2.5%	SPOT	[14]
Railroad Valley Playa, US	3%	SPOT, AVHRR and LANDSAT TM	[15]
Railroad Valley Playa, US	4.31%	MERIS	[6]
Western Queensland, Australia	1.2 to 5.2%	LANDSAT ETM+ & TM	[11]
Saharan & Arabian Deserts	1 to 2%	METEOSAT	[28]

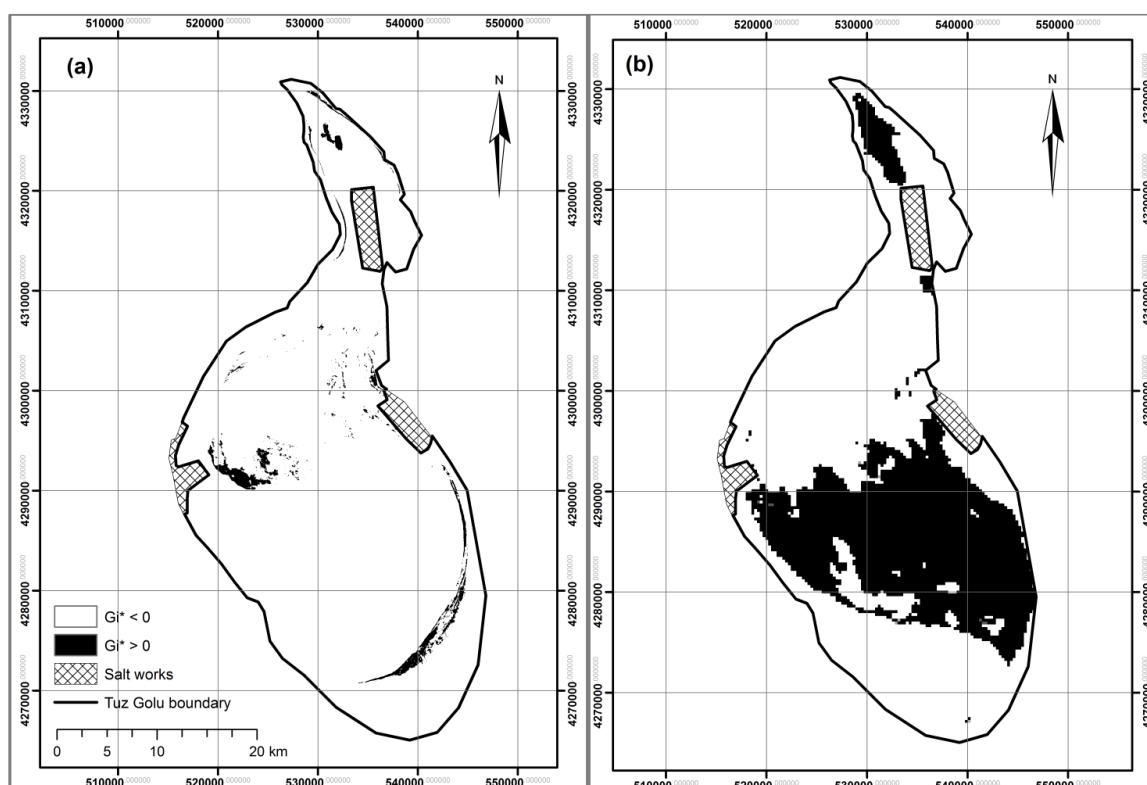
The results showed that there was a consistently spatially homogeneous area located at the mid central part of Tuz Gölü year after year (Figure 9). However, much of the remaining area within Tuz Gölü (approximately 97%) is more variable and not all of it is stable year after year. This is natural, considering that Tuz Gölü is a seasonal salt lake that experiences intra-annual and inter-annual variations in the surface as well as fluctuating water levels. These variations could be responsible for the modification of the surface structure properties. The authors of [16] investigated the water coverage of the lake as it changes in a year. The findings showed that, the lower parts of the lake remain covered in water for up to six and eight months in a year which is longer than the middle and upper parts of the lake which are only covered by water for five months in a year. This means that, the upper middle parts remain dry in most of the years and suitable for vicarious calibration campaigns.

The spatially homogeneous areas computed within a local window may vary from year to year for the same local window. This necessitates the analysis of a series of multi-temporal imagery to clearly inform on the temporal stability of a surface using both CV and Gi^* . Spatial homogeneity clearly has a temporal dimension since there may be areas, which progressively shift from being homogenous to heterogeneous and vice versa from year to year [27]. The temporal dynamics of this vary from year to year and may explain why the area with $Gi^* > 0$ was much more restricted for some years than others. Ancillary data would be required to assess this in more detail, but such data were not available for this study.

A legitimate criticism of the approach presented in this paper is that the issue of temporal stability is only partially addressed. The methodology identified homogenous areas (*i.e.*, relatively bright clusters with low variance) at a single point in time. The time series was then queried to find areas that were

persistently homogeneous over time. However, this ignored the possibility that the mean HDRF of a persistently-homogeneous area may go up and down over time. Indeed, a similar drawback was noted for spatial-CV (*i.e.*, adjacent areas could have the same CV but different means). This mean component of temporal variability is not addressed in this paper. Indeed, this issue may remain unresolved owing to the lack of data required to characterize the atmosphere on the specific dates of the images used. Without such data the mean component will be inaccurate; however, the variance component, which is key for the CV and G_i^* computations, will remain unchanged.

Figure 11. (a) Spatial uniformity map of current study based on Landsat TM for the month of August for 1984, 1989 (2 images), 1998, 2000, 2003, 2006, 2009 (3 images) integrating bands 2, 3, 4; (b) Spatial uniformity map based on MODIS satellite images of July and August (2004–2006) using only red and near infra-red bands (adapted from [16]).



The spatial homogeneity of the surface is scale-dependent and wavelength dependent, and the persistence of homogeneous areas over time depends upon the time scale of the observations. This is shown by comparing the results of this study with those of [4]. Their approach was based on applying only the Getis statistics to MODIS Red and Near Infra-Red bands with 250 m pixel size acquired in the month of July–August 2004 and July–August 2006. It led to a spatial homogeneity map quite different from that found in the present study (Figure 11). This shows that a site which is homogeneous over the spatial support (pixel size) measured by a coarse resolution sensor, and stable over short time periods, may be much more heterogeneous and dynamic when studied in more detail and over a longer time period.

Finally, there is another dimension to spatial scale that has also not been considered in this paper. That relates to the choice of the 3×3 window, which was justified in Sections 3.2 and 3.3. The window size determines the area within which the autocorrelation is assessed. The authors of [17] and [27]

analyzed the variability of the G_i^* statistic for a range of window sizes. They argued that this gave a fuller understanding of the spatial structure and extent of clustering across their study areas. This knowledge would be useful because the eventually-chosen calibration target will have an area that is larger than 3×3 pixels and must also be sampled in the field. This study has adopted the 3×3 window, in line with previous recommendations in the literature. The analysis was repeated using a 9×9 window, leading to similar results. Nevertheless, future research should be directed at a more comprehensive understanding of the spatial structure of vicarious calibration targets. This should include analysis of the global statistics (e.g., the variogram) as well as local clustering at a range of window sizes.

6. Conclusion

In this study the coefficient of variation, CV, was used to identify areas with low local variability, within a 3×3 window. This is in line with previous studies. However, the CV alone was insufficient for identifying homogenous areas. It is not a spatial measure. Furthermore, adjacent areas may have low CV, but quite different means. Hence the Getis statistic, G_i^* , a local indicator of spatial association (LISA), was used to identify bright clusters.

The resulting methodology used the G_i^* and CV to identify bright clusters ($HDRF > 30\%$) with low local variability. Both the CV and G_i^* were computed within 3×3 windows, in line with earlier studies. Temporally stable areas were those that showed these properties consistently for all Landsat TM images analyzed over a 25 year period. The final identified site was a contiguous area of at least $720 \text{ m} \times 330 \text{ m}$ which met the aforementioned criteria consistently over time.

Retrieved HDRF from the proposed VC site on Tuz Gölü showed uncertainty, as measured by the CV, in the range of 0.99% to 3.99% for Landsat TM bands 2, 3 and 4. The results fit well within the range of other reported uncertainties of calibration test sites in the world. This showed the potential of Tuz Gölü as a calibration target for EO activities, as well as the effectiveness of the proposed methodology.

Acknowledgements

Vincent Odongo was supported by the EU Erasmus Mundus scholarship programme. Robin Wilson (University of Southampton) is thanked for providing his RTW tools IDLTM routines and Dr. Joanna Nield (University of Southampton) is thanked for sharing her insights on the geomorphology and dynamics of desert surfaces and playas. We appreciate the comments of four anonymous reviewers whose comments helped improve this manuscript.

Author Contributions

Vincent O. Odongo performed the analysis and wrote the first draft of the paper. Nicholas A. S. Hamm supervised the analysis and provided extensive feedback and contributions to the written paper. Edward J. Milton instigated the research and contributed to the written paper.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Bannari, A.; Omari, K.; Teillet, P.M.; Fedosejevs, G. Potential of Getis statistics to characterize the radiometric uniformity and stability of test sites used for the calibration of Earth observation sensors. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2918–2925.
2. Fox, N.P. Validated Data and Removal of Bias through Traceability to SI Units. In *Post-Launch Calibration of Satellite Sensors*; Morain, S.A., Budge, A.M., Eds.; Taylor and Francis: London, UK, 2004.
3. Teillet, P.M.; Horler, D.N.H.; O'Neill, N.T. Calibration, validation, and quality assurance in remote sensing: A new paradigm. *Can. J. Remote Sens.* **1997**, *23*, 401–414.
4. Gürol, S.; Behnert, I.; Özen, H.; Deadman, A.; Fox, N.; Leloğlu, U.M. Tuz Gölü New CEOS reference standard test site for infrared visible optical sensors. *Can. J. Remote Sens.* **2010**, *36*, 553–565.
5. Leloğlu, U.M.; Gürbüz, H.; Özen, H.; Gürol, S. Characterization of the Spatial Uniformity of the Tuz Gölü Calibration Test Site. In Proceedings of ISPRS Workshop on Modeling of Optical Airborne and Spaceborne Sensors, Istanbul, Turkey, 11–13 October 2010.
6. Kneubühler, M.; Schaepman, M.E.; Thome, K. Long-Term Vicarious Calibration Efforts of MERIS at Railroad Valley playa (Nevada)—An update. In Proceedings of the Second Working Meeting on MERIS and AATSR Calibration and Geophysical Validation (MAVT-2006), Frascati, Italy, 20–24 March 2006.
7. Teillet, P.M.; Barsi, J.A.; Chander, G.; Thome, K.J. Prime candidate Earth targets for the post-launch radiometric calibration of space-based optical imaging instruments. *Proc. SPIE* **2007**, doi:10.1117/12.733156.
8. Thome, K.J. Absolute radiometric calibration of Landsat 7 ETM+ using the reflectance-based method. *Remote Sens. Environ.* **2001**, *78*, 27–38.
9. Schaepman-Strub, G.; Schaepman, M.E.; Painter, T.H.; Dangel, S.; Martonchik, J.V. Reflectance quantities in optical remote sensing—Definitions and case studies. *Remote Sens. Environ.* **2006**, *103*, 27–42.
10. Milton, E.J.; Schaepman, M.E.; Anderson, K.; Kneubühler, M.; Fox, N. Progress in field spectroscopy. *Remote Sens. Environ.* **2009**, *113*(S1), S92–S109.
11. de Vries, C.; Danaher, T.; Denham, R.; Scarth, P.; Phinn, S. An operational radiometric calibration procedure for the Landsat sensors based on pseudo-invariant target sites. *Remote Sens. Environ.* **2007**, *107*, 414–429.
12. Woodcock, C.E.; Strahler, A.H. The factor of scale in remote sensing. *Remote Sens. Environ.* **1987**, *21*, 311–332.
13. Rondeaux, G.; Steven, M.D.; Clark, J.A.; Mackay, G. La Crau: A European test site for remote sensing validation. *Int. J. Remote Sens.* **1998**, *19*, 2775–2788.
14. Scott, K.P.; Thome, K.J.; Brownlee, M.R. Evaluation of Railroad Valley playa for use in vicarious calibration. *Proc. SPIE* **1996**, doi:10.1117/12.256090.
15. Bannari, A.; Omari, K.; Teillet, P.M.; Fedosejevs, G. Multi-sensor and multi-scale survey and characterization for radiometric spatial uniformity and temporal stability of Railroad Valley playa (Nevada) test site used for optical sensor calibration. *Proc. SPIE* **2004**, doi:10.1117/12.506864.

16. Gurol, S.; Ozen, H.; Leloglu, U.M.; Tunali, E. Tuz Gölü New absolute radiometric calibration test site. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *37*, 35–39. Available Online: http://www.isprs.org/proceedings/XXXVII/congress/1_pdf/06.pdf (accessed on 12 March 2014).
17. Getis, A.; Ord, J.K. The analysis of spatial association by use of distance statistics. *Geogr. Anal.* **1992**, *24*, 189–206.
18. Getis, A. Spatial Dependence and Heterogeneity and Proximal Databases. In *Spatial Analysis and GIS*; Taylor and Francis; London, UK, 1994; pp. 105–120.
19. Ord, J.K.; Getis, A. Local spatial autocorrelation statistics: Distributional issues and an application. *Geogr. Anal.* **1995**, *27*, 286–306.
20. Gokmen, M.; Vekerdy, Z.; Lubczynski, M.W.; Timmermans, J.; Batelaan, O.; Verhoef, W. Assessing groundwater storage changes using remote sensing-based evapotranspiration and precipitation at a large semiarid basin scale. *J. Hydrometeorol.* **2013**, *14*, 1733–1753.
21. Camur, M.Z.; Mutlu, H. Major-ion geochemistry and mineralogy of the salt lake (Tuz Gölü) basin, Turkey. *Chem. Geol.* **1996**, *127*, 313–329.
22. Kilic, O.; Kilic, A.M. Salt crust mineralogy and geochemical evolution of the salt lake (Tuz Gölü), Turkey. *Sci. Res. Essays* **2010**, *5*, 1317–1324.
23. Helder, D.L.; Markham, B.L.; Thome, K.J.; Barsi, J.A.; Chander, G.; Malla, R. Updated radiometric calibration for the Landsat-5 thematic mapper reflective bands. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 3309–3325.
24. Richter, R. *Atmospheric/Topographic Correction for Satellite Imagery—ATCOR-2/3 User Guide*; Version 6.3; Remote Sensing Data Centre Publication DLR-IB 565-01/07; German Aerospace Centre: Wessling, Germany, 2007.
25. Wilson, R.T. *RTWTools*; 2011. Available online: <http://www.rtwtools.rtwilson.com/>. (accessed on 14 March 2014).
26. Wulder, M.; Boots, B. Local spatial autocorrelation characteristics of remotely sensed imagery assessed with the Getis statistic. *Int. J. Remote Sens.* **1998**, *19*, 2223–2231.
27. LeDrew, E.F.; Holden, H.; Wulder, M.A.; Derksen, C.; Newman, C. A spatial statistical operator applied to multirate satellite imagery for identification of coral reef stress. *Remote Sens. Environ.* **2004**, *91*, 271–279.
28. Cosnefroy, H.; Leroy, M.; Briottet, X. Selection and characterization of Saharan and Arabian desert sites for the calibration of optical satellite sensors. *Remote Sens. Environ.* **1996**, *58*, 101–114.
29. Gu, X.; Guyot, G.; Verbrugghe, M. Analysis of the spatial variability of a test-site: The example of La Crau. *Photo Interpret.* **1990**, *90*, 39–52. (In France)
30. Thomas, D.S.G. *Arid Zone Geomorphology: Process, Form and Change in Drylands*, 3rd ed.; Wiley: Chichester, UK, 2011.
31. Magee, J.W. Late quaternary lacustrine, groundwater, aeolian and pedogenic gypsum in the Prungle Lakes, southeastern Australia. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **1991**, *84*, 3–42.