Indian Journal of Science and Technology Vol. 4 No. 6 (June 2011) ISSN: 0974- 6846

Preparation of digital data sets on land use/land cover, soil and digital elevation model for temperature modelling using Remote Sensing and GIS techniques

Sailesh Samanta[1], Dilip Kumar Pal[1], Debasish Lohar[2] and Babita Pal[1]
[1]*Department of Surveying and Land Studies, The PNG University of Technology, Morobe Province, Lae, 411, Papua New Guinea;*
[2]*Department of Physics, Atmospheric Science Research Group, Jadavpur University, West Bengal, India*
rsgis.sailesh@gmail.com; dkpal_2000@yahoo.com; dlohar_atmos@yahoo.com; mundey2@gmail.com

Abstract

Remote Sensing (RS) and Geographic Information Systems (GIS) are becoming powerful tools in climatological modelling. This study proposes an empirical methodology to prepare digital data set of land use/land cover, soil and digital elevation model (DEM) using RS and GIS techniques. The study area is Gangetic West Bengal and its neighborhood in the eastern India, where a number of weather systems occur throughout the year. Gangetic West Bengal is a region of strong heterogeneous surface with several weather disturbances. Standard false color composite (Std FCC) bands (green, red and near Infrared bands) of LANDSAT-7, ETM+ sensor are used to produce land use/land cover dataset. DEM is built using the contours and spot heights which are collected from Topographical maps of the region. With the help of soil region map of West Bengal and soil region map of India the soil texture dataset is built. All the data sets are converted to the spatial resolution of 1km and 5km, which can be used as the independent variables for climatological modelling (specifically air temperature modelling) of the study area.

Keywords: Remote Sensing, Geographical Information System, Land use/land cover, Soil, Relief.

Introduction

Surface cover is very important for governing air temperature variations of a particular area (Eliasson & Svensson, 2003). Surface properties, such as texture, colour, roughness, moisture regime and the type & density of vegetation canopy play the dominant role in determining surface atmosphere energy exchanges. In most cases, the observed air temperatures varies significantly according to ambient land cover types, and through latent heat transfer the air temperatures tend to decrease as the amount of vegetated area around the sites increases. These land cover effects are slightly stronger at night than during the day, and they weakened as the amount of cloud cover or the wind speed increases, especially at night (Yokobori & Ohta, 2009). The urban area increases the absorption and storage of incoming solar radiation because the canyon geometry of an urban area increases the surface area and leads to multiple reflections. The situation impedes exit avenues of the radiation culminating in higher absorption and thereby leading to enhanced temperature regime vis-à-vis rural settings. The characteristic of urban surfaces is manifested in reduced evapo-transpiration thus augmenting sensible heating. In addition, urban areas are sources of anthropogenic heat (Oke, 1982). As a result, the energy balance in an urban area is disturbed. This effect of the urban climate on air temperatures is called urban heat island (UHI), which refers to the temperature difference between urban areas and surrounding cool rural areas (Yokobori & Ohta, 2009). Vegetation can play an important role in moderating this phenomenon, because the incoming radiation energy in a vegetated area is converted into latent heat during the day and the sensible heat flux is lower than that of a built-up area covered by asphalt or concrete (Jonsson, 2004). The shading of vegetation reduces the radiant absorption of the ground. At night, a vegetated area with short plants or an area of open bare ground may cool rapidly due to its higher radiative cooling rate (Spronken-Smith & Oke 1999). Thus a multitude of factors are responsible for air temperature differences. As land cover directly affects the energy balance characteristics, it is essential to understand the air temperature differences in the study area. The air temperature is also influenced by soil temperature which in turn is a function of primarily soil organic matter, moisture regime and the mineral content (cumulative effect of which gives rise to the soil colour). Average monthly soil temperature near the soil surface is calculated using equations developed by Parton (1984). These equations calculate maximum soil temperature as a function of the maximum air temperature and the canopy biomass (lower for high biomass) while the minimum soil temperature is a function of the minimum air temperature and canopy biomass (higher for higher biomass). The actual soil temperature used for decomposition of soil organic matter and plant growth rate function is the average of the minimum and maximum soil temperatures. Changes in soil temperature will lag behind the changes in air temperature if the difference in air temperature between months is greater than 2°C (Eliasson and Svensson 2006).

In "the atmosphere" we understand that air temperature usually decreases with an increase in elevation through the troposphere. The decrease in temperature with elevation is called the environmental lapse rate of temperature or normal lapse rate of temperature. The normal lapse rate of temperature is the average lapse rate of temperature of 0.65^{o} C / 100 meters. The environmental lapse rate of temperature is the actual vertical change in temperature on any given day and can be greater or less than 0.65^{o} C / 100 meters. Also recall that the decrease in temperature with height is caused by increasing distance vertically from the datum – the mean sea level where the highest concentration of Green House Gases (GHGs) exist that drop drastically with height as the heating of air takes place predominantly from below and not from above.

Study area and environmental conditions

The study area includes a geographical area extending from 21° N to 25° N and 85° E to 89° E. Parts of the eastern India and western part of Bangladesh come under the above longitudinal and latitudinal extension. The study area includes Gangetic West Bengal, eastern part of Jharkhand, north-eastern part of Orissa and a small southern part of Bihar in India. The region consists in a semi-arid part in the west and south-west and relatively more moist and cool in the south and the east. Being the tail end of Chhotonagpur plateau, the terrain of the semi-arid part is more complex with higher elevation and permanent vegetation compared to the south and the eastern part of the study area. The region experiences four seasons having typical weather systems: monsoon (June through September) season, characterized by south-west monsoonal flow having a monsoon trough line through the region causing most of the annual rainfall; post-monsoon season (October and November), being the transition season between the monsoonal flow and westerlies of winter season (December through February) in the upper levels, region experiences tropical storm on occasions; a part of the region on the other hand experiences the impact of western disturbances in winter; and frequent thunderstorm activity in the pre-monsoon season (March through May), which gives temporary relief from the scorching heat of summer. Therefore, throughout the year the study region experiences synoptic scale disturbances, such as monsoons and mesoscale weather systems leading to more complex patterns of climatic variables. This emphasizes the need for a better procedure to map the climatic variables in absence of dense network of observations.

Materials used

Different types of data sets are used for preparation of digital data sets of land use/land cover, soil and DEM within the area. These are (i) One of the most widely used DEM data sources is the elevation information provided by the Shuttle Radar Topography Mission (SRTM) (Coltelli *et al*., 1996), As with most other DEM sources, the SRTM data requires significant levels of pre-processing to ensure that there are no spurious artifacts in the data such as pits, spikes and patches of no data (Dowding *et al*., 2004) that would cause problems in later analysis. In the case of the SRTM data, these patches of no data are filled, preferably with auxiliary sources of DEM data, like-topographical maps. Both SRTM data sets and topographical map are used for this study. (ii) West Bengal soils sheet of National Bureau of Soil Survey (NBSS) and soil region map of National Atlas & Thematic Mapping Organization (NATMO) with different group of soil were used to generate soil texture map. (iii) Optical bands with Std FCC of LANDSAT 7 ETM+ satellite images were used to find out the land use/land cover classes in the study area. All other details of the variables are given in the Table 1.

Methodology

Total one hundred and five (105) land use/ land cover samples are identified with their associated features in the field. About 75% of the ground truth blocks are used for training sets and the rest 25% are used for assessment of the accuracy of results (Pal *et al*., 1991). Maximum likelihood algorithm under supervised classification technique is used to generate land use/land cover of the area from the satellite image. With the help of ERDAS IMAGINE s/w recode is performed on the classified product. Accuracy assessment (Table 3 & 4) is performed using the ground truth on the classified image.

Soil texture classification is performed using two referenced soil maps, West Bengal soil sheets 1 to 4 of National Bureau of Soil Survey and Land Use Planning (NBSS & LUP) national atlas of India, soil region of National Atlas & Thematic Mapping Organization (NATMO), Department of Science and Technology, Government of India, Kolkata, India. All the soils are categorized into nine (9) soil texture groups according to their characteristics. To generate the soil texture map of the study area, ArcInfo and ERDAS IMAGINE s/w are used.

The original SRTM DEM is used to produce contours or points. Processing is made on a void by void basis. In cases when a higher resolution auxiliary DEM is available, point coverage is produced of the elevation values at the centre of each cell of the auxiliary DEM within void areas. When no high resolution auxiliary DEM is available, the 30 second SRTM30 DEM is used as an auxiliary for large voids. For areas with a high resolution auxiliary DEM: The contours and points surrounding the hole and inside the hole are interpolated to produce a hydrologically sound DEM using the TOPOGRID algorithm in Arc/Info. TOPOGRID is based upon the established algorithms of Hutchinson (1988; 1989), designed to use contour data (and stream and point data if available) to produce hydrologically sound DEMs. This process interpolates through the no-data holes, producing a smooth elevation surface where no data was originally found.

On the other hand, DEM is also generated from topographical maps. Single map rectification is performed using the geographical coordinate system and WGS 84 datum with a RMS error of 0.02. ERDAS IMAGINE 9.0 is used to rectify topographical maps of the study area using the geographical Coordinate system and WGS 84 datum with a RMS error of 0.25. The entire study area does not appear in single sheet. All the geo-referenced topographical map sheets of the area are mosaicked to get the entire study area. Using the similar projection as well as datum for all the soil maps are mosaicked in the

Indian Journal of Science and Technology Vol. 4 No. 6 (June 2011) ISSN: 0974- 6846

Table1. Different variables used for this study

Sl. No	Different Variables	Scale/Resolution	Year/ Range	Source
1	Topographical map	1:250000	1960	University of Texas Libraries, Austin
2		1:50000	1973-1980	Survey of India, Kolkata
3	LANDSAT-7, ETM+;	30 m	2000 - 2001	University of Maryland Institute for Advanced Computer Studies
4	National Atlas of India, Soil Region	1:2000000	1981	National Atlas & thematic Mapping Organisation, Department of Science and Technology, Govt. of India
5	Shuttle Radar Topography Mission (SRTM) data	3-arc seconds	2003	*ftp://e0srp01u.ecs.nasa.gv*

Table 2. RGB color presentations and statistics for different land use/ land cover classes

Recode value	Land use/ land cover	RGB Color presentation	Area (in sq km)	% of area
0	Water (Sea or inland)	0.000:0.000:1.000	19782	10.93
1	Marshy land	0.250:0.880: 0.820	2818	1.56
2	Forest	0.000:0.390:0.000	23048	12.73
3	Shrub land (Mixed or open)	1.000:0.750:0.800	22078	12.19
4	Crop land (Agricultural land)	1.000:1.000:0.120	112348	62.04
5	Built up area (Settlement or industries)	0.440:0.040:0.200	989	0.55

Table 3. Classification error matrix report for land use/land cover (pixel Size- 30m x 30m)

Classified Data	Water	Marshy	Forest	Shrub	Agriculture	Urban & built-up
Water	5	0	0	0	0	0
Marshy	0	4	0	0	0	0
Forest	0	0	6	0	0	0
Shrub	0	0	1	5	1	0
Agriculture	0	0	1	2	22	0
Urban & built-up	0	0	0	0	0	3
Column Total	5	4	8	7	23	3

Table 4. Classification accuracy totals report for land use/land cover (pixel Size- 30m x 30m)

Class Name	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy
Water	5	5	5	100.00%	100.00%
Marshy	4	4	4	100.00%	100.00%
Forest	8	6	6	75.00%	100.00%
Shrub	7	7	5	71.43%	71.43%
Agriculture	23	25	22	95.65%	88.00%
Urban & built-up	3	3	3	100.00%	100.00%
Totals	50	50	45		

Overall Classification Accuracy = 90.00%

Table 5. RGB color presentations and statistics for different soil texture classes

Row No.	Soil texture classes name	Recode Value	RGB Color presentation	Area (in sq km)	% of area
1	Water	0	0.000:0.000:1.000	16422	9.1
2	Loamy sand	1	1.000: 1.000: 0.880	11467	6.3
3	Sandy loam	2	0.930:0.510: 0.930	37742	20.8
4	Silty loam	3	1.000:0.840:0.000	1302	0.7
5	Loamy	4	0.750:0.750:0.750	4360	2.4
6	Sandy clay loam	5	0.690:0.190:0.380	54248	30.0
7	Silty clay loam	6	0.650:0.160:0.160	27662	15.3
8	Clay loam	7	0.500:1.000: 0.000	1557	0.9
9	Sandy clay	8	0.000:0.390:0.000	5238	2.9
10	Silty clay	9	1.000:1.000:0.000	21065	11.6

Indian Journal of Science and Technology Vol. 4 No. 6 (June 2011) ISSN: 0974- 6846

Fig.1. Land use/land cover of the study area, generated from LANDSAT-7, ETM+ satellite image

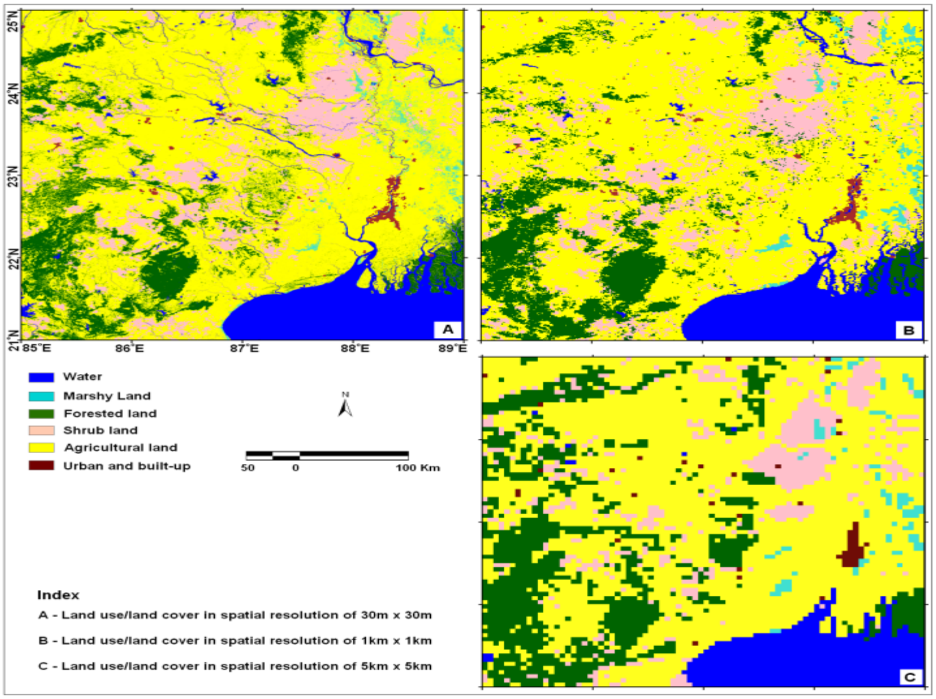
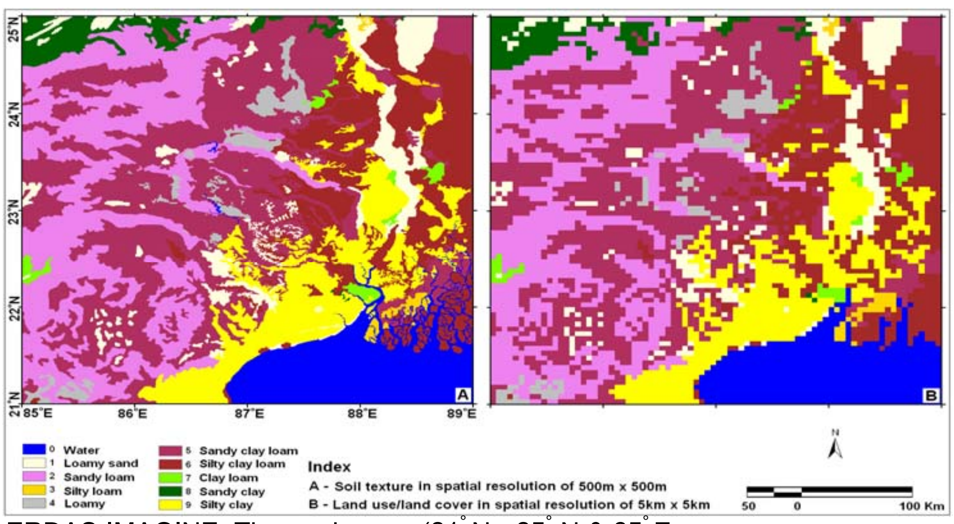


Fig.2. Soil texture of the study area, generated from soil region map and NBSS soil sheet



ERDAS IMAGINE. The study area (21° N - 25° N & 85° E - 89° E) extracted by sub-setting from the mosaic map using the coordinates of four corners. To generate the digital elevation model of the study area ArcInfo and ERDAS IMAGINE s/w are used. Another point layer is generated in the same s/w environment and the spot heights are collected according to the longitudinal and latitudinal value. Then spot height attribute for each sample location is introduced to the point layer. Arc coverage, which represents the contours, is used to build 3D surfacing in the ERDAS IMAGINE s/w environment. All point data are interpolated to the raster through krigging interpolation techniques in the ArcGIS s/w environment. Prepared data sets in this process are compared with the SRTM data set to find out the accuracy of both data sets.

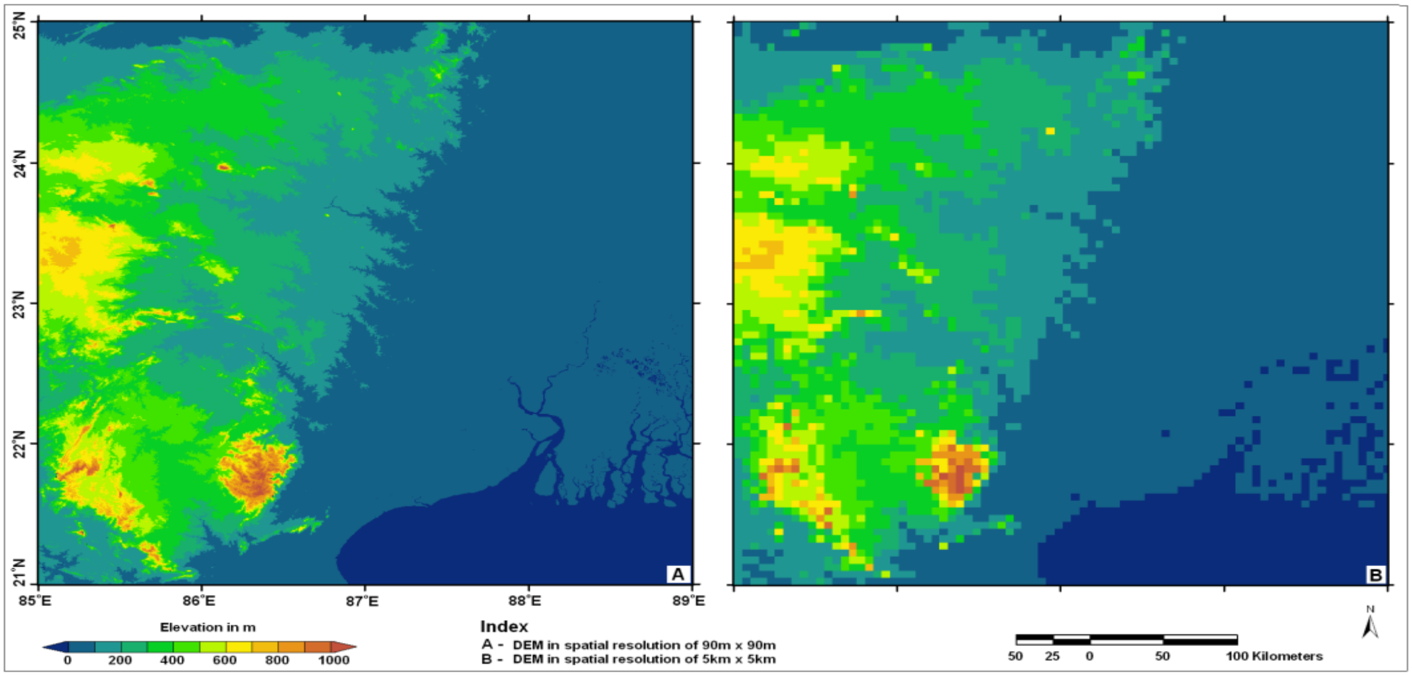
Results

Using ERDAS IMAGINE and ArcGIS software, different raster layers were generated in different spatial resolution. The land use /land cover data set were generated from the digital image classification of LANDSAT, ETM+ satellite images. This classification was performed taking six classes (Table 2). These were - (i) Water (Sea or inland water bodies), (ii) Marshy land, (iii) Forest, (iv) Shrub land (Mixed or open), (v) Crop land (Agricultural land) and (vi) built up area (Settlement or industries). The western part was relatively more heterogeneous having forested area, crop land and shrub land etc. as shown in Fig.1. Therefore the climatic variables patterns were expected to be more complex. From the tonal values of original Std FCC imageries, land cover classes were selected and for classification a signature editor was created to supervise the computer software. About 90% of overall accuracy was obtained from 30m x 30m classified image, whereas only 84% overall accuracy was obtained from 1km x 1km data set. After striking a satisfied level of accuracy (Table 3 & 4), the map composition was attempted for final outputs in form of map. Three sets of data/map were generated with different spatial resolutions to represent the land use/land cover in various details, viz. of 30m, 1km and 5 km spatial resolution.

Soil texture data set was generated by the process of digitization from soil region maps of the area. West Bengal soil sheets of NBSS and soil region map of NATMO were used for this purpose. Soil types were categorized into nine (9) soil texture groups (Table 5) according to their characteristics, namely- loamy sand, sandy loam, silty loam, loamy sandy, clay loam, silty clay loam, clay loam, sandy clay, and silty clay. The southern part was relatively more heterogeneous than western and northern parts as shown in Fig.2. Two sets of data/map were generated to represent the soil texture with first one of 500m and second one of 5km spatial resolution.

Indian Journal of Science and Technology Vol. 4 No. 6 (June 2011) ISSN: 0974- 6846

Fig.3. Digital elevation of the study area, generated from SRTM data and topographical maps



SRTM data were used to generate DEM for the study area. We had to use topographical maps in order to fill the data gaps. Using the 3-D analysis process the digital elevation model was generated from contours, which had been digitized from topographical maps. The entire data range was classified in to 11 groups with the interval of 100 m altitude. Two sets of data/map were generated to represent the digital elevation model having 90m and 5km spatial resolutions, as shown in Fig.3. It showed that the elevation of the eastern and southern parts of the study area had been less than 100m while the northern part had been relatively more elevated (>1000m).

Same methodology and materials were used to prepare digital data sets of land use/land cover (Fig.4), soil texture (Fig.5) and DEM (Fig.6) in 1km spatial resolution for entire West Bengal. An attempt was made using simple equation (Equation-1) to find out average distribution of seasonal temperature with the help of these three independent parameters.

$$L_\beta + S_\gamma + E_\delta = T \quad \text{(1)}$$

[Where, T= Predicted temperature, L_β= Predicted temperature weight by land use, S_γ= Predicted temperature weight by soil type and E_δ= Predicted temperature weight by elevation]

To find out the predicted temperature weight by land use, soil and elevation following three equations (equation 2, 3 & 4) were solved.

$$L_\beta = \frac{\sum_{x,i} \frac{t_i x}{M}}{N} \quad \text{(2)}$$

$$S_\gamma = \frac{\sum_{y,i} \frac{t_i y}{M}}{N} \quad \text{(3)}$$

$$E_\delta = \frac{\sum_{z,i} \frac{t_i z}{M}}{N} \quad \text{(4)}$$

[Where, $\beta \propto x, i$; $\gamma \propto y, i$; $\delta \propto z, i$; t= Observed temperature; x= Weight of land use/ land cover (33% to 40%); y= Weight of soil texture (12% to 20%); z= Weight of elevation (30% to 40%); i= Post-monsoon season; M= x + y + z; N= Number of treatment sample points]

Different weight values were selected for different parameters and for each category of data set in model. Taking the example of land use, there were different types of land use / land cover classes and as such separate weight values were assigned for each and every class. A model calculation has been made to choose different weight for the parameters. Mean temperature data set of India Meteorological Department (IMD) was used as one input in order to derive the equation. In the output post-monsoon temperature distribution raster layer was generated (Fig.7). Result showed the clear effect of DEM in the post-monsoon temperature distribution map. The northern part of West Bengal was elevated more than 1500m as shown in the DEM and as a result the temperature was very low in that part of the study area. Not only DEM, the spatial variation of temperature also depended on land use/land cover and soil properties as shown in the Fig.7.

Indian Journal of Science and Technology Vol. 4 No. 6 (June 2011) ISSN: 0974- 6846

Fig.4. Land use/land cover of West Bengal (1km pixel size)

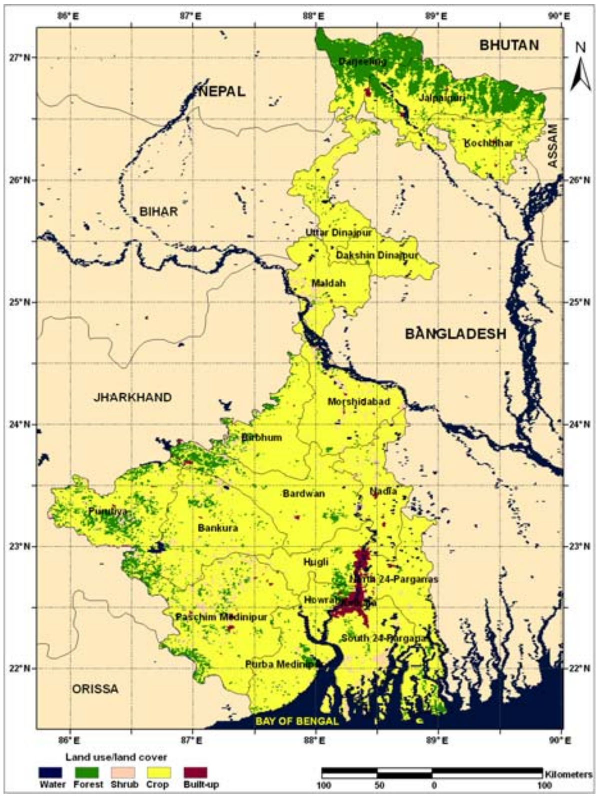


Fig.5. Soil texture of West Bengal (1km pixel size)

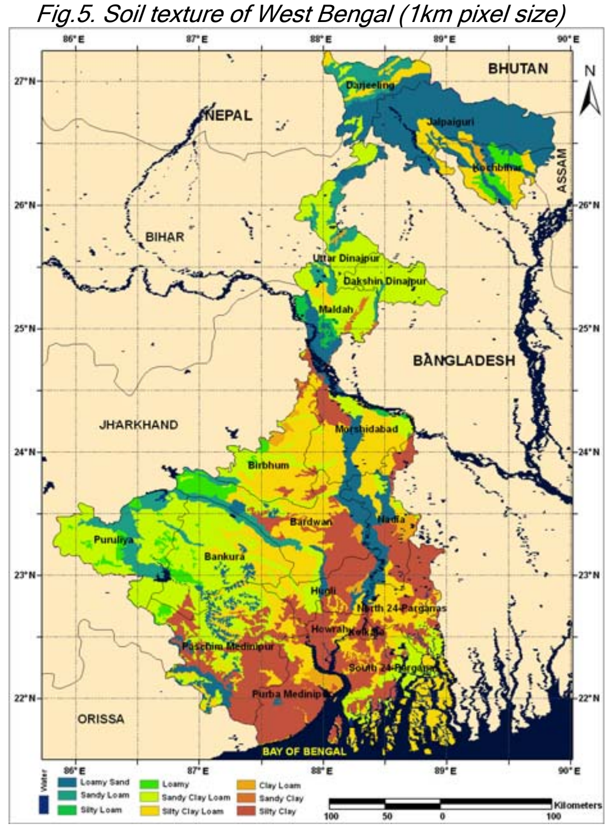


Fig.6. Digital elevation of West Bengal (1km pixel size)

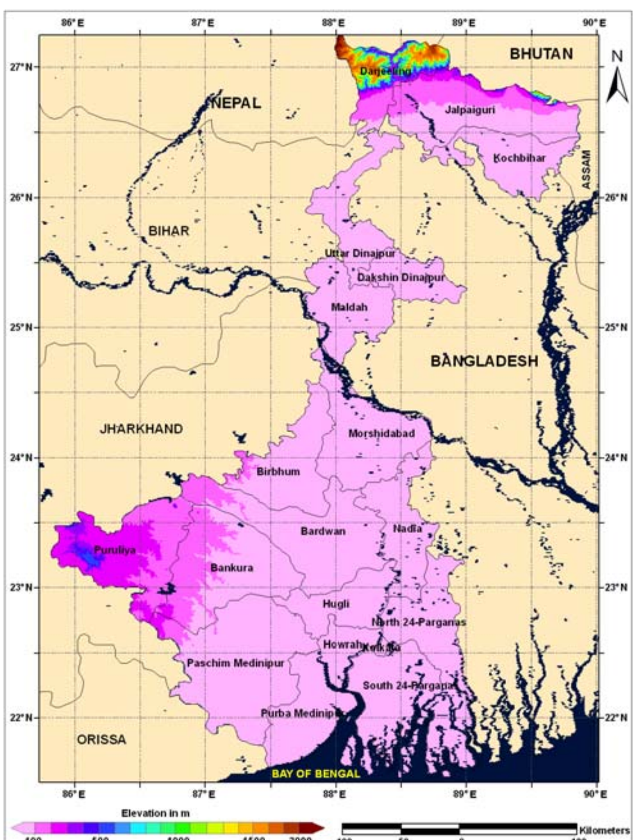
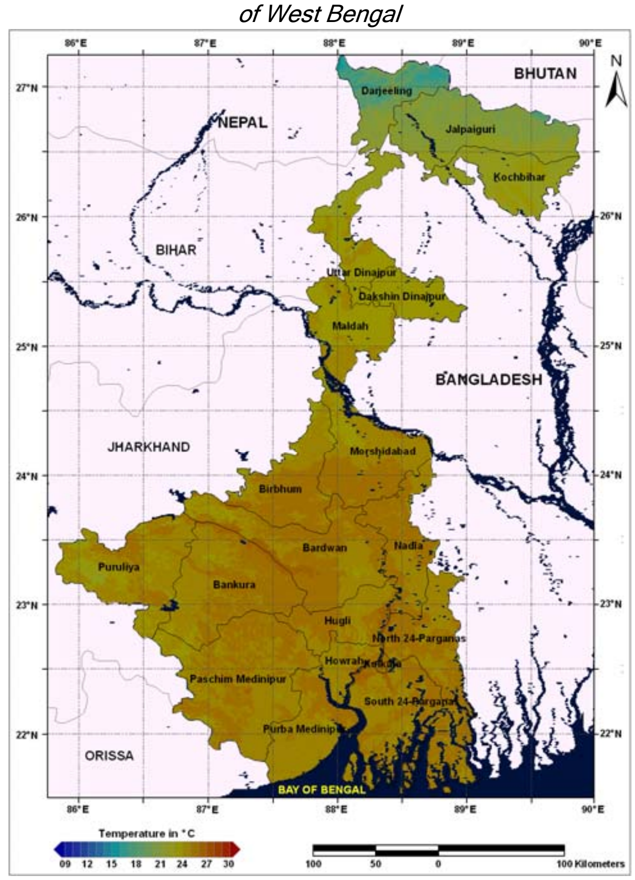


Fig.7. Post-monsoon temperature distribution of West Bengal



Conclusion

Daily temperature data are relatively continuous and can often be estimated relatively accurately using data from only a small number of nearby stations and using a small number of environmental factors as covariates. For example, after the inclusion of elevation as one covariate, including further covariates yielded only minor further overall improvements in the accuracy of estimations, although additional factors such as distance from the coast, distance from urban areas and local shelter did improve estimations locally for certain types of weather system (Jarvis & Stuart, 2001).

In the atmosphere it is understood that air temperature usually decreases with an increase in elevation through the troposphere. The air temperature is also influenced by the soil temperature which owes its existence to factors like organic matter content, moisture regime and the mineral content. The spatial and temporal distribution of temperature from day to day has a major impact on the agriculture, environment and economy of many countries. Spatially distributed estimates of temperature are required as inputs to many environmental models. ANNs can be used to estimate daily temperature with similar accuracy to the most sophisticated of the conventional methods, whilst requiring less stringent assumptions to be met about the data conditions (Rigol *et al.*, 2001). So land use/land covers (Fig.1), soil textures (Fig.2) and altitude/elevation (Fig.3) can be used for modeling of daily or monthly temperature as independent variables along with dependent variables of the weather and climate. Evidently, other variables like distance from the sea, latitudinal location, continentality, solar radiation, permanent wind directions (trade wind, westerlies) and cloudiness factor, aspect etc also can be used as additional parameters to improve the model while amplifying the complexity in the process. This model can be run for calculating the mean monthly and seasonal temperature for other areas as well, provided the weightage of the parameters are recalculated for relevant areas.

Acknowledgement

One of the authors (SS) expresses sincere gratitude to Department of Surveying and Land Studies Papua New Guinea University of Technology for providing digital image interpretation laboratory facility to carry out the research work. The authors remain also grateful to the anonymous referees for their valuable comments and suggestions.

References

1. Coltelli M, Fornaro G, Franceschetti G, Lanari R, Migiaccio M, Moreira JR, Papathanassaou KP, Puglisi G, Riccio D and Schwabisch M (1996) SIR-C/X-SAR multifrequency multipass interferometry: A new tool for geological interpretation. *J. Geophys. Res*. 101, 23127-48.
2. Dowding S, Kuuskivi T and LI X (2004) Void fill of SRTM elevation data – principles, processes and performance. In: Images to Decisions: Remote Sensing Foundations for GIS Applications, ASPRS, Fall Conf., Sep. 12-16, Kansas City, MO, USA.
3. Eliasson I and Svensson MK (2006) Spatial air temperature variations and urban land use- a statistical approach. *Meteorol. Appl*. 10 (2), 135-149.
4. Hutchinson M (1988) Calculation of hydrologically sound digital elevation models. Third Intl. Sym. on Spatial Data Handling, Columbus, Ohio, Intl. Geographical Union.
5. Hutchinson M (1989) A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *J. Hydrol.* 106, 211-232.
6. Jarvis CH and Stuart N (2001) A comparison between strategies for interpolating maximum and minimum daily air temperatures. b. The interaction between guiding variable and interpolation method. *J. Appl. Meteorol*. 40, 1075-1084.
7. Jonsson P (2004) Vegetation as an urban climate control in the subtropical city of Gaborone, Botswana. *Intl. J. Climatol.* 24, 1307-1322.
8. Oke TR (1982) The energetic basis of the urban heat island. *Quart J. R. Met. Soc*. 108, 1-24.
9. Pal DK, Venkataramana I, Sudhakar S and Krishnan N (1992), Land use / Land cover mapping using IRS-1A data. In: Natural Resources Management - a new perspective. National Natural Resources Management System, Bangalore. pp: 360-366.
10. Parton WJ (1984) Predicting soil temperatures in a short grass steppe. *Soil Sci*. 138, 93-101.
11. Rigol JP, Jarvis CH and Stuart N (2001) Artificial neural networks as a tool for spatial interpolation. *Intl. J. Geographical Information Sci*. 15(4), 323-343.
12. Spronken-Smith RA and Oke TR (1999) Scale modelling of nocturnal cooling in urban parks. *Boundary-Layer Meteorol.* 93, 287-312.
13. Yokobori T and Ohta S (2009) Effect of land cover on air temperatures involved in the development of an intra-urban heat island. *Climate Res*. 39, 61-73.