

## Mesoscale model compatible IRS-P6 AWiFS-derived land use/land cover of Indian region

Biswadip Gharai<sup>1,\*</sup>, P. V. N. Rao<sup>2</sup> and C. B. S. Dutt<sup>3</sup>

<sup>1</sup>Atmospheric Chemistry and Processes Studies Division, Earth and Climate Science Area, and

<sup>2</sup>Remote Sensing Area, National Remote Sensing Centre, Indian Space Research Organisation, Balanagar, Hyderabad 500 037, India

<sup>3</sup>Karnataka State Rural Development and Panchyat Raj University, Gadag 582 101, India

**Mesoscale models, in general, are run using the US Geological Survey (USGS) 25-category land use/land cover (LU/LC) data available at different spatial resolutions. The USGS data over the Indian region suffers from two types of errors, viz. misclassification of LU/LC data and non-availability of up-to-date satellite-based LU/LC data. To improve the accuracy and capture interannual changes better, the LU/LC data generated by the National Remote Sensing Centre (NRSC) using IRS-P6 AWiFS with 56 m basic resolution have been scaled to 5, 2 min and 30 sec resolution which is available at yearly intervals. In the next step, the Indian region of USGS data was replaced with IRS-P6 AWiFS-derived data and made compatible to MM5 and WRF mesoscale models. Thus the resultant product is a global USGS LU/LC data with the Indian region replaced by the information originally derived from AWiFS 56 m resolution imagery, for the years 2004–05 to 2012–13 (nine cycles). This communication describes the required LU/LC data format for MM5 and WRF models and the methodology adopted for compatible product generation. In addition, accuracy of AWiFS-derived LU/LC data converted to 30 sec resolution has also been determined. The present effort will provide the necessary reference for the atmospheric modelling community to address the Indian satellite based model compatible LU/LC data product. These data products are currently available on Bhuvan, the NRSC/ISRO geospatial portal.**

**Keywords:** Land use/land cover data, land-surface processes, mesoscale model, spatial resolution.

LAND use/land cover (LU/LC) changes are considered to be one of the most important factors affecting the regional climate and thus become an area of public concern. LU/LC inputs are a critical part of the meteorological modelling system. The role of the land surface is particularly important in driving boundary layer evolution and ultimately precipitation patterns. Inaccurate LU/LC information often leads to large errors in surface energy

---

\*For correspondence. (e-mail: g.biswadip@gmail.com)

fluxes and thus errors in boundary layer states. In meteorological models, many land-surface variables like leaf area index (LAI), fractional vegetation cover, canopy height, emissivity, albedo, surface roughness, etc. are commonly defined as a function of LU/LC via a 'lookup table'. These vegetation-related variables exert significant control on the surface temperature energy balance and subsequently on boundary layer processes and states, most importantly, moisture and temperature profiles. The lookup table approach assumes a one-to-one relationship between the surface variable and LU/LC category, with no variability represented within a LU/LC category. In many model applications, seasonal or monthly parameter values are defined, providing an annual cycle of vegetation phenology. In some applications, satellite observations are used to define a subset of these variables, primarily albedo, LAI and fractional vegetation cover.

Changes in surface parameters from semiarid land conditions to cultivated lands tend to consistently increase the potential for moist convection during daytime heating hours, which cause evolution of the boundary layer height structure and growth of convective available potential energy (CAPE) as well as area-averaged rainfall<sup>1</sup>. The impact of continental-scale land-cover change (LCC) was explored on the Australian climate<sup>2</sup>. The literature<sup>3</sup> shows that the reduced surface roughness following LCC largely explains the simulated changes in rainfall by increasing moisture divergence over southwestern Australia and increasing moisture convergence inland.

The 25-category (including no data class) US Geological Survey (USGS) LU/LC dataset (USGS, 1994) being used in most meteorological modelling/air quality applications, is deficient in its characterization of the urban/suburban landscape<sup>4,5</sup>. This Global Land Cover Characterization (GLCC) LU/LC dataset was derived from 1 km AVHRR data spanning April 1992 through March 1993 (modified level 2)<sup>6</sup>. Modellers often use outdated USGS GLCC 30 sec (~1 km) LC data. The release of the 2001 National Land Cover Data (NLCD) products at 30 m cell resolution for the US and 2001 NASA MODIS LC data at 1 km cell resolution for the globe have created lot of interest among modelling community. Ran *et al.*<sup>7</sup> found improvement by incorporating 2001 NLCD data in the WRF model.

Two types of errors exist with the default WRF LU/LC (USGS) data, viz. (1) up-to-date satellite data are not used for LU/LC layer creation, and (2) LU/LC data include misclassification<sup>8</sup>. In a vast country like India with varied LC, it is more important to utilize up-to-date LU/LC data in the mesoscale model. Dutta *et al.*<sup>9</sup> studied the impact of vegetation dataset derived from SPOT satellite by ISRO (Indian Space Research Organisation) versus the USGS vegetation dataset on the simulation of the Indian summer monsoon. In their study, multi-date SPOT vegetation data (spatial resolution ~1 km) for the year 2000 were re-aggregated/regrouped to USGS 25

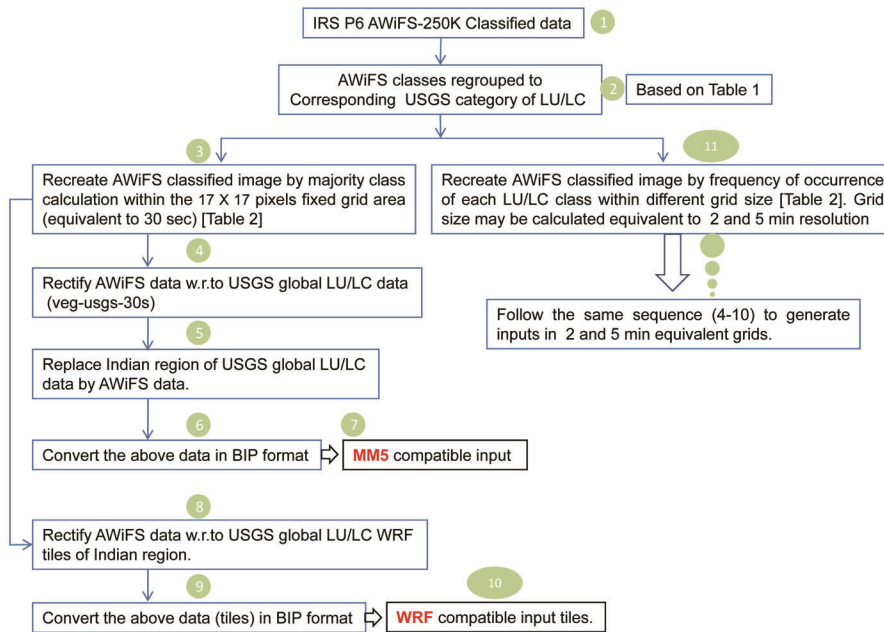
classes, spatially aggregated to 10 min grid size and used as input to the MM5 model.

LU/LC is one of the dynamic features that undergoes rapid changes due to anthropogenic and associated developmental activities and is amenable to satellite remote sensing. Hence up-to-date information on LU/LC is one of the important requisites for regional climate models as the land surfaces influence partitioning of incident solar radiation to different energy fluxes. The meteorological variables of air temperature, humidity and wind speed at the surface are influenced by the surface energy fluxes. It has been observed that there are errors in the LU/LC classes over the Indian region in the USGS dataset, especially in urban, water bodies and forested areas. Hence there is a need to include accurate and up-to-date satellite-based LU/LC data as input to MM5/WRF model.

With the above background, the objective of this study was to generate WRF/MM5 model compatible AWiFS LU/LC datasets for numerical weather prediction/climate modelling of the Indian region. Prior to the present study there was no effort to make Indian satellite-based LU/LC data compatible with MM5 and WRF models. This communication describes the methodology adopted for MM5/WRF model compatible AWiFS-based LU/LC product generation. The present effort will provide the necessary reference for the modelling community to address the Indian satellite based model compatible LU/LC data product.

The Natural Resources Census (NRC) programme was taken up as part of the Natural Resources Repository (NRR) activity under the National Natural Resources Management System (NNRMS) of the Department of Space (DOS), Government of India (GoI) to map different natural resources in the country. The National Land Use Land Cover mapping project was taken up during 2004–05 with an objective to undertake 'rapid assessment of national-level land use land cover on 1 : 250,000 scale using multi-temporal IRS-P6 AWiFS datasets'. The study has been completed for nine crop calendar years from 2004–05 to 2012–13. AWiFS is one of the sensors on-board IRS-P6 satellite having four spectral bands with 56 m ground resolution and five days revisit.

Multi-temporal IRS-P6 AWiFS data covering *kharif* (August–November), *Rabi* (January–March) and *Zaid* (April–May) crop seasons were used to address spatial and temporal variability in cropping pattern and other LCCs. The datasets were supplemented with sensor data from IRS WiFS/Terra-Aqua MODIS over cloud and quality-affected AWiFS datasets. The thematic classification of AWiFS data involves geo-referencing of multi-temporal datasets with Lambert conformal conic projection and WGS 84 datum. Further, AWiFS datasets were converted to top of the atmosphere reflectance data to minimize sun angle variations between the dates of acquisition. The detailed methodology of IRS-P6 AWiFS LU/LC classification, area statistics and accuracy is described in detail



**Figure 1.** Methodology to generate IRS-P6 AWiFS land use/land cover datasets in WRF/MM5 model compatible format.

in the NRSC Bhuvan website (<http://bhuvan.nrsc.gov.in/gis/thematic/tools/document/LULC250/1213.pdf>). The LU/LC maps generated were field-verified for assessing classification accuracy. The USGS LU/LC data were derived from MODIS 1 km spatial resolution data, whereas the Indian effort was based on IRS-P6 AWiFS 56 m resolution data. In general, mapping at high resolution and upscaling to coarser resolution bring better details than directly mapping at coarser resolution. Moreover, in AWiFS-classified data, use of ground truth information resulted in realizing high accuracy.

Figure 1 shows the methodology to generate AWiFS LU/LC datasets in WRF/MM5 model compatible format. By default, MM5/WRF mesoscale models use USGS-generated regular latitude/longitude inputs on vegetation (LU/LC), terrain elevation, land–water mask, soil types, vegetation fraction and deep soil temperature. It has 24 categories (plus one no-data category), whereas LU/LC data derived from Indian satellite IRS-P6 AWiFS has 19 categories. Since the current objective is to make AWiFS-classified data compatible to USGS global data for wider use by modellers, AWiFS (from Indian satellite) classified data were first regrouped into USGS 24 categories based on Table 1. In this table, the second column corresponds to USGS 24 LU/LC classes; the fourth column represents IRS-P6 AWiFS-derived 19 classes and the seventh column shows class ID of the AWiFS classes (from column 3) that need to be recoded for the purpose of matching with USGS LU/LC data as given in column 6 (recoded AWiFS classes). Evergreen forest class ID = 13 (in column 5) within the Indian Himalayan region is recoded to ‘evergreen Needle leaf (class ID = 14)’ as

given in Table 1. Then recoded AWiFS data (in \*.img format) in ERDAS imagine platform (version: 2014; Intergraph Corporation, Part of Hexagon, USA) are exported to BIP (band interleaved by pixel) format for further processing.

The USGS LU/LC global data are assumed to be valid at the centre of a grid box. Hence there are  $360 \times 180$  data points for  $1^\circ$  global data;  $((360 \times 2) \times (180 \times 2))$  for 30 min;  $((360 \times 120) \times (180 \times 120))$  data points for 30 sec data (Table 2). As shown in Table 2, higher resolution 30 sec data are equivalent to 0.925 km ( $\sim 1$  km) and coarse resolution data of 2 and 5 min are equivalent to 3.7 and 9.25 km spatial resolution respectively.

For the 30 sec compatible data generation from AWiFS classified data, majority classes are calculated based on a new fixed grid, accommodating  $17 \times 17$  AWiFS pixels (Table 2), and then replacing the Indian region of global USGS LU/LC data (of 24 classes) by IRS-P6 AWiFS-derived LU/LC data. The global data generated are then exported to BIP format to produce MM5 compatible output equivalent to 30 sec. The exported data ‘veg-usgs.30s’ are global USGS LU/LC data with Indian region replaced by AWiFS-derived data.

Similarly, 30 sec AWiFS majority filtered data were geo-corrected with respect to USGS WRF tiles (24 classes) and these data were used to replace USGS global WRF tiles of the Indian region; these tiles were exported to BIP format to generate WRF compatible tiles of the Indian region. Total number of 30 sec WRF tiles was 16.

For coarse resolution data of 2 and 5 min, new grids were created accommodating 66 and 165 AWiFS pixels respectively (Table 2). Then frequency of occurrence of

**Table 1.** Regrouping criteria to match IRS-P6 AWiFS-based land use/land cover (LU/LC) data with USGS LU/LC classes

CLS-ID	USGS	CLS-ID	Actual AWiFS class	CLS-ID	Recoded-AWiFS	Recode AWiFS
1	Urban and built-up La	1	Build up	1	Settlement	1
2	Dryland crop past	2	<i>Kharif</i> only	2	<i>Kharif</i> crop only	2
3	Irrigation crop past	3	<i>Rabi</i> only	3	<i>Rabi</i> crop/ <i>zaid</i> /double/triple crop	3 + 4 + 5
4	Mixed dryland/irrigation	4	<i>Zaid</i> only	4		
5	Crop/grs. Mosaic	5	Double/triple	5	Current fallow	6
6	Crop/wood mosaic	6	Current fallow	6		
7	Grassland	7	Plantation/orchard	7	Grassland	12
8	Shrubland	8	Evergreen forest	8	Scrub/deg. Forest/scrubland	10 + 15
9	Mix shrub/Grs	9	Deciduous forest	9		
10	Savanna	10	Scrub/Deg. Forest	10		
11	Deciduous broadleaf forest	11	Littoral swamp	11	Deciduous forest	9
12	Deciduous needle leaf	12	Grassland	12		
13	Evergreen broadleaf forest	13	Other wasteland	13	Evergreen forest	8
14	Evergreen needle leaf	14	Guilied	14	Evergreen forest (Himalayan region)	8 after rerode
15	Mixed forest	15	Scrubland	15	Plantations/orchard/shifting cultivation	7 + 18
16	Water bodies	16	Water bodies	16	Water bodies	16
17	Herbaceous wetland	17	Snow covered	17		
18	Wooded wetland	18	Shifting cultivation	18	Littoral swamp (mangrove)	11
19	Barren or sparsely veg.	19	Rann	19	Other wasteland/guillied/ran	13 + 14 + 19
20	Herbaceous tundra			20		
21	Wooded tundra			21		
22	Mixed tundra			22		
23	Bare ground tundra			23		
24	Snow or ice			24	Snow cover	17
25	No data			25		

**Table 2.** Different grid sizes and number of AWiFS pixels in each grid.

Grid (long. × lat. wise)	Total number of grids	Grid resolution	Grid resolution (kilometres)	Number of AWiFS pixels in each grid
360 × 12 × 180 × 12	9,331,200	5 min	9.25	165
360 × 30 × 180 × 30	58,320,000	2 min	3.7	66
360 × 120 × 180 × 120	933,120,000	30 sec	0.925	17

each of the LU/LC classes in IRS-P6 AWiFS data at different resolutions mentioned above was calculated. Figure 2 shows AWiFS LU/LC frequency data of India with 2 min resolution. AWiFS-derived 2 min equivalent frequency image created as mentioned above has 24 classes. Also, a single layer with no data is stacked to the frequency image to produce an image of 25 classes, as required to make USGS-equivalent data. Then AWiFS-derived LU/LC (25 classes) image was geo-corrected with respect to the global USGS 2 min LU/LC data. USGS global LU/LC (2 min) data of the Indian region were then replaced with geo-corrected AWiFS LU/LC 2 min data; this generic file was exported to BIP format to produce MM5 input ‘veg-usgs.02’, which are global USGS LU/LC data with the Indian region replaced by AWiFS-derived data.

AWiFS LU/LC frequency data (24 classes) were geo-corrected with respect to USGS 2 min resolution WRF LU/LC tiles mosaic data and the global USGS LU/LC (2 min) data of Indian region replaced by geo-corrected AWiFS data. The USGS 2 min WRF LU/LC individual tiles of the Indian region (nine numbers) were used to

extract tile-wise AWiFS replaced 2 min data. Then nine individual AWiFS-replaced tiles were exported to BSQ format to produce nine WRF compatible AWiFS data.

In a similar manner, 5 min resolution AWiFS-derived data were made compatible to MM5 and WRF models. MM5/WRF compatible file naming convention was in accordance with USGS global data. Figure 3 shows the comparative display of USGS and AWiFS LU/LC data. All individual classes in USGS LU/LC were randomly checked with ground information collected while verifying AWiFS classified data. Large omissions in water bodies and urban area were found as evident in the figure, which could be mainly due to the use of relatively old satellite database in USGS LU/LC. In the vegetation classification, gross error in vegetation types was also observed. Similar types of errors were also reported by other researchers<sup>8</sup>.

While geo-correcting AWiFS data with respect to USGS (Indian region) LU/LC data, common ground control points (GCPs) were identified on both the classified data. The coefficients for two coordinate transformation equations were compared based on second-order

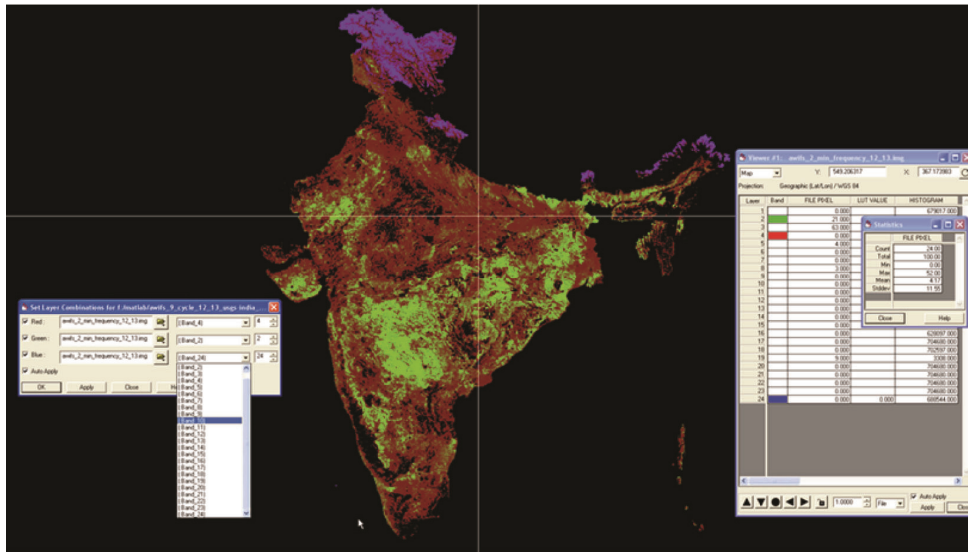


Figure 2. IRS-P6 AWiFS-based LU/LC frequency data for 2 min resolution.

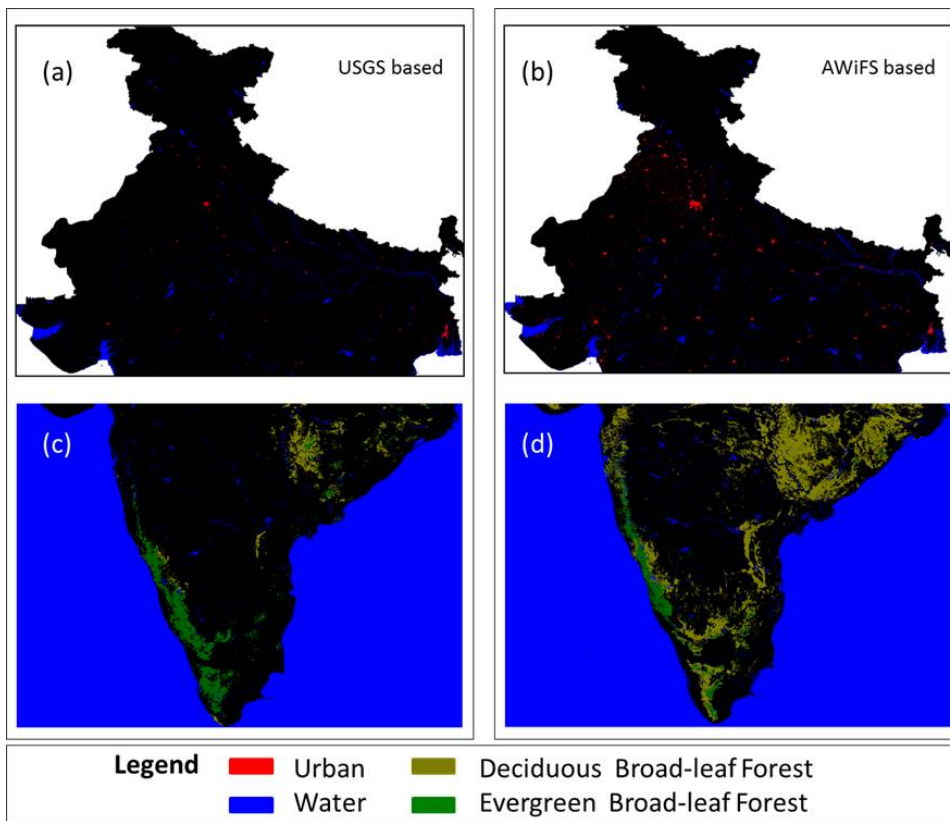


Figure 3. Comparative display of USGS and AWiFS LU/LC data (part of India) for 30 sec resolution, where many classes are misrepresented in the USGS LU/LC data. *a*, USGS urban; *b*, AWiFS urban; *c*, USGS forest; *d*, AWiFS forest.

polynomial regression between GCPs on the USGS and AWiFS classified data. Alternate GCPs were generated till the root mean square error was less than 0.5 pixels size and then nearest neighbourhood resampling was applied to obtain geo-corrected data. At every stage of geo-correction ‘\*.gcc’ files were saved and used for output generation of different cycles to maintain consistency in data production. Each time after producing BIP/BSQ

format compatible to MM5/WRF, the data were checked by importing into ‘\*.img’ format for visualization. As an example, MM5 2 min input ‘veg-usgs.02’ data were imported into BIP format with row=5400, column=10800, 25 classes and displayed in ERDAS imagine. It was ensured that while replacing the Indian region with AWiFS-derived LU/LC data, rest of the global data were not contaminated. This was done by subtracting Indian region

data (AWiFS) from the AWiFS-replaced USGS global data to get zero fill in the Indian region. The AWiFS-replaced global data at 30 sec, 2 and 5 min resolution were made available to different modelling groups in India, including the National Centre for Medium Range Weather Forecasting, Ministry of Earth Sciences, GoI; Satish Dhawan Space Centre, Sriharikota, ISRO, GoI, and various participating institutions in the Continental Tropical Convergence Zone programme conducted by the Department of Science and Technology, GoI.

USGS LU/LC data were derived from 1 km spatial resolution MODIS data whereas Indian satellite based LU/LC data were prepared based on 56 m basic spatial resolution of AWiFS data and subsequently these were made different spatial resolutions as mentioned in the text. The 30 sec IRS-P6 AWiFS-derived LU/LC data, which were regrouped with respect to USGS classified (30 sec) data, are considered to estimate their accuracy. A total of 320 stratified random points were generated using ERDAS imagine to estimate the accuracy. As it has been observed visually that urban, water bodies and forest areas are mostly misrepresented in USGS classification, more number of random points was allocated accordingly. For error calculation we used GPS-aided ground-truth data, already available classified data over India and the Google Maps ([www.google.co.in](http://www.google.co.in)) as the reference. The number of points generated for different classes, commission and omission errors; user and producer accuracy; overall accuracy and overall kappa statistics were generated. A total of 320 points were generated, of which urban, deciduous broadleaf, evergreen broadleaf and water bodies were represented by 50, 30, 40 and 60 points respectively. The assessment showed an overall accuracy from the stratified random sampling method for the ninth cycle AWiFS-derived LU/LC data with 30 sec resolution to be 90.94%, with an overall kappa statistic 0.90.

The initial results from the model (WRF) run using IRS-P6 AWiFS-derived LU/LC data suggest considerable improvement in the prediction of meteorological parameters. As an example, Unnikrishnan *et al.*<sup>10</sup> used 'Unified Model coupled with Joint UK Land Environment Simulator land surface model' to study the impact of AWiFS-derived LU/LC data of 2012–13 on weather prediction. They reported that both wet and dry weather case studies of prediction showed improvement in forecast by incorporating the ISRO LU/LC (AWiFS-derived) data. They also reported an improved prediction of regional rainfall pattern using AWiFS-derived LU/LC data.

At present, model compatible AWiFS LU/LC data are available in 30 sec, 2 and 5 min resolution for modellers through NRSC Bhuvan/NICES geospatial portal (<http://www.nrsc.gov.in/nices>).

The present study detailed the procedure adopted to make the IRS-P6 AWiFS-based LU/LC data compatible with global USGS LU/LC data for the Indian region at different resolutions. (i) The MM5/WRF compatible

AWiFS-derived LU/LC inputs have the potential to considerably improve prediction of meteorological parameters. (ii) The outcome of the present study will provide necessary reference for the modelling community to address the Indian satellite based model compatible LU/LC data product.

1. Nataly, P. and Pinhas, A., Effect of land use modification on potential increase of convection: a numerical mesoscale study over south Israel. *J. Geophys. Res.*, 2001, **106**(D19), 621–634.
2. Narisma, G. T. and Pitman, A. J., The impact of 200-years land cover change on the Australian near-surface climate. *J. Hydro-meteorol.*, 2003, **4**, 424–436.
3. Pitman, A. J., Narisma, G. T., Pielke Sr, R. A. and Holbrook, N. J., Impact of land cover change on the climate of southwest Western Australia. *J. Geophys. Res.*, 2004, **109**, D18109; doi:10.1029/2003JD004347 1-12.
4. Quattrochi, D. A., Lapenta, W. M., Crosson, W. L., Estes, M. G., Limaye, A. and Khan, M., The application of satellite-derived, high-resolution land use/land cover data to improve urban air quality model forecasts, NASA/TP-2006-214710, November 2006.
5. Byun, D. W., Kim, S. T., Czader, B. and Cheng, B., Modeling effects of land use and land cover modifications on the urban heat island phenomena and air quality in Houston. Project H17 Final Report prepared for Houston Advanced Research Center, Houston, TX, USA, 2004.
6. Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L. and Merchant, J. W., Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *Int. J. Remote Sensing*, 2000, **21**(6&7), 1303–1330.
7. Ran, L., Pleim, J. and Gilliam, R., Impact of high resolution landuse data in meteorology and air quality modeling system. In *Air Pollution Modelling and its Application* (eds Steyn, D. G. and Rao, S. T.), Springer Science+Business Media BV, 2010, XX; doi:10.1007/978-90-48-3812-8.
8. Sertel, E., Robock, A. and Ormecci, C., Impacts of land cover data quality on regional climate simulations. *Int. J. Clim.*, 2009; doi:10.1002/JOC.2036.
9. Dutta, K. S., Das, S., Kar, S. C., Mohanty, U. C. and Joshi, P. C., Impact of vegetation on the simulation of seasonal monsoon rainfall over the Indian subcontinent using a regional model. *J. Earth Syst. Sci.*, 2009, **118**(5), 413–440.
10. Unnikrishnan, C. K., Gharai, B., Mohandas, S., Mamgain, A., Rajagopal, E. N., Gopal, R. I. and Rao, P. V. N., Recent change on land use/land cover over Indian region and its impact on the weather prediction using Unified model. *Atmosp. Sci. Lett.*, 2016, doi:10.1002/asl.658.

ACKNOWLEDGEMENTS. We thank Dr V. K. Dadwal, former Director, NRSC for the necessary support to host the mesoscale compatible AWiFS derived LU/LC data in NRSC Bhuvan geospatial portal. We also thank Project Director and all the team members of the 'National land use and land cover mapping using multi-temporal AWiFS data (LULC-AWiFS, 250 K)' project. Dr P. S. Roy, former Dy. Director, RS&GIS, NRSC is gratefully acknowledged for his help in regrouping AWiFS classified data equivalent to USGS data. Late Dr K. V. S. Badarinath is sincerely acknowledged for initiating the present work. B.G. thanks Mr Sivaprasad Reddy, Mr M. Y. Aslam, Mr R. V. N. Srinivas and Mr M. Naresh Kumar for their help.

*Authors' contribution.* B.G. has conceived and designed the study, carried out the study work and wrote the manuscript; P.V.N.R. has reviewed the manuscript and worked for hosting the model compatible LU/LC data in NRSC Bhuvan geospatial portal; C.B.S.D. is responsible for identifying AWiFS LU/LC classes which are needed to be regrouped with respect to USGS classes.

Received 12 August 2016; revised accepted 18 September 2018

doi: 10.18520/cs/v115/i12/2301-2306