MODIS TIME SERIES REMOTE SENSING FOR EPIDEMIOLOGICAL MODELLING

Markus Neteler

ITC-irst, SSI/MPBA
Via Sommarive, 13
38050 Povo (Trento), Italy
Email: neteler *at* itc *dot* it

ABSTRACT

This paper reports on the processing of time series of MODIS satellite data in a Geographical Information System (GIS). The data preparations for the GIS usage are described with focus on the reprojection from MODIS/SIN to national coordinate systems and the application of the MODIS quality maps. We explain subsequent filtering of Land Surface Temperature maps with an outlier detector to eliminate originally undetected cloud pixels. Further analysis of time series is briefly discussed as well as related landscape epidemiological applications in the field of tick-bourne diseases.

1 INTRODUCTION

In epidemiological modelling, survey data are usually collected at sampling sites and then regionalised in Geographical Information Systems (GIS). To enhance the spatial data density, continuous field data such as land surface temperatures (LST), snow coverage, and vegetation indices are commonly derived from satellite data. The launches of the new satellite systems *Terra* (December 1999) and *Aqua* (May 2002) significantly improve the situation of data availability for scientific purposes and predictive epidemiological studies. The *Moderate Resolution Imaging Spectroradiometer* (MODIS) is a key instrument on both Terra and Aqua satellites. As they deliver daily two global coverages at 250m (Red, NIR), 500m (MIR) and 1000m resolution (TIR), they are most interesting to support epidemiological studies. Usually one week after acquisition the data are available to the public.

The orbit of Terra around the Earth is scheduled to pass from north to south across the equator in the morning, while Aqua passes with reverted direction from south to north over the equator in the afternoon. Terra, crossing the equator at about 10:30 a.m. local solar time, is in a sun-synchronous orbit with a delay of 30 minutes with respect to LANDSAT-7. The further orbital parameters are equal to those of LANDSAT-7. MODIS is a whisk-broom sensor with 36 channels ranging from visible to thermal-infrared (GSFC/NASA, 2003). Data are delivered at 250m (2 channels), 500m (5 channels) and 1000m resolution (29 channels).

MODIS can be considered as a much enhanced successor of the AVHRR instrument onboard the NOAA series of satellites. MODIS improves upon the performance of AVHRR by providing both higher spatial resolution and greater spectral resolution. This paper focuses on two of the numerous MODIS data products: Land Surface Temperatures (LST), and Vegetation Index 16-day composites (NDVI and EVI).

2 EPIDEMIOLOGICAL RISK MAPS AND STUDY REGION

The methods described in this paper have been implemented to extend an epidemiological study about the exposure risk to Lyme disease transmitted by ticks in the Autonomous Province of Trento, Italian Alps. The predictions are carried out through the analysis of the distribution of *Ixodes ricinus* (L.) nymphs infected with *Borrelia burgdorferi* s.l. with a model based on tree-based classifiers. This model is supported by a Geographical Information System (Merler et al., 1996; Rizzoli et al., 2002; Furlanello et al., 2003). In this project data on *I. ricinus* (L.) density, assessed by dragging the vegetation in 438 sites during 1996 were cross-correlated with the digital cartography of a GIS. The area includes the Autonomous Provinces of Bolzano and Trento and the Province of Belluno (Italy), a region of approximately 18,000 km^2 area size. The complex terrain has an elevation range from slightly above sea level to 3800 m (Bolzano: elev. range 3700m, mean elev. 1800m, mean slope 26°; Trentino: elev. range 3700m, mean elev. 1400m, mean slope 26°; Belluno: elev. range 3200m, mean elev. 1500m, mean slope 27°).

The density of available meteorological stations is highly varying and often concentrates in the valleys. To improve the risk mapping at high resolution, other sources of climatic and vegetational data are required. This demand can be fulfilled by using remote sensing data. The integration of satellite data into epidemiological research enhances the spatio-temporal resolution of climatological data in particular in mountainous regions where climatic stations and ground surveys are unavailable or sparse (Hess et al., 2002).

3 MODIS DATA PREPROCESSING FOR GIS USAGE

MODIS data sets are delivered as "Base Level Swath Data" as well as "gridded data". Grid data are projected in either Integerised Sinusoidal (ISIN, Level V003) or Sinusoidal (SIN, Level V004) projections. Both data types require further preprocessing before they can be used in a GIS. Especially ISIN is usually unsupported in most of "off the shelf" and free GIS and image processing software. The "MODIS Reprojection Tool" (MRT, U.S. Geological Survey 2004) can be used to reproject ISIN or SIN to a more common projection (e.g. UTM) or to a national grid systems such as Gauss-Boaga (Italy). Furthermore MRT allows for geographical subsetting, writes the output to standard data formats such as GeoTIFF, and is executable on various operating systems. From April 2004 the V003 product is removed from the USGS archives as V004 becomes available. The V004 data quality is significantly improved, especially for processing inland water pixels (Wan, 2003).

Further pre-processing steps comprise the pixel-wise application of the quality map provided along each data product and an outlier detection for certain MODIS products to minimise the presence of low quality pixels. This is explained in greater detail in the following part.

3.1 Preprocessing of Land Surface Temperatures (LST) data

MODIS Land Surface Temperature and Emissivity (LST/E) products are mapping land surface temperatures and emissivity values. The underlying algorithms use other MODIS data as input, including geolocation, radiance, cloud masking, atmospheric temperature, water vapour, snow, and land cover. Temperatures are extracted in Kelvin; accuracy of 1 Kelvin is yielded for materials with known emissivities (Wan, 1999).

After reprojection, we are applying pixel-wise the related quality maps to the LST maps. This step is required due to limitations in the official cloud detection algorithm used to create the land surface temperature quality maps. In particular the cloud detection during the night pass is error prone at cloud corners. To overcome this problem, a simple outlier detection for the minimum temperatures has been integrated into the proposed procedure:

$$lower\_boundary = 1st\_Quartile - 1.5 * (3rd\_Quartile - 1st\_Quartile) \quad (1)$$

These values have to be determined with respect to the pixel-wise varying data usability. These variations occur due to aerosol and cloud presence and other quality limitations as indicated in the quality map. We used only LST maps with a sufficient number of pixels in this calculation. To avoid that a few pixels influence the overall outlier statistics, we have chosen a minimum availability of at least 25% valid LST pixels. The LST maps are filtered then on a monthly base with these mean lower boundary thresholds. Each threshold is calculated from the monthly mean of the lower boundary values which are available twice a day (day and night pass of MODIS). The proposed simple outlier filter aims at removing all pixels which contain cloud top surface temperatures remaining undetected from the NASA algorithm.

3.2 Preprocessing of Vegetation Indices (NDVI/EVI)

The two MODIS vegetation indices, the well known Normalised Difference Vegetation Index (NDVI) and a new developed Enhanced Vegetation Index (EVI) are spectral measures of the amount of vegetation present on the ground. The MODIS-NDVI describes the relative "greenness" of the Earth's vegetation on a scale of minus one (-1) to plus one (+1) and is intended to continue the more than 20 years of NOAA/AVHRR-derived NDVI. The magnitude of NDVI is related to the level of photosynthetic activity in the observed vegetation. The EVI is MODIS-specific and offers improved sensitivity in high biomass regions while being less sensitive to atmospheric aerosol scattering (especially smoke from burning vegetation). It also minimises the influence of background interference caused by bare soil reflecting off the ground (Huete et al., 1999).

These vegetation indices can be integrated into epidemiological models to reflect vegetation dynamics. The 16-day composite product minimises cloud cover problems by substituting a cloud covered pixel with a later uncontaminated pixel within a 16-days period. These composites reflect the current vegetation status at a sufficient temporal resolution.

In order to use these MODIS map products in epidemiological applications, the beforehand reprojected NDVI/ EVI maps have to be filtered according to the related quality bit pattern maps. These quality maps contain information about pixel quality, aerosol contents, cloud, water or snow presence etc. and they are encoded in 16 bit. A bit pattern module was developed for GRASS GIS to perform this operation (`r.bitpattern`).

4 MODIS TIME SERIES ANALYSIS

The study has been carried out with GRASS GIS software (Neteler and Mitasova, 2004). The recent implementation of general time series processing (`r.series`) for GRASS raster maps supports univariate statistics for a series of MODIS scenes. By selecting various time ranges and operators, a number of indicators can be calculated.

4.1 Validation and applications of daily LST data

To validate the usability of MODIS/Terra data in epidemiological studies as an enhancement of data availability, the monthly mean temperatures of selected meteorological stations and the related MODIS data at the same coordinates have been investigated. Fig. 1 shows monthly mean temperatures of Cavedine meteorological station (Trentino, Italy) compared to mean LST results from MODIS/Terra. Despite missing pixels due to cloud and high aerosol presence the

Table 1: Number of observations for calculation of monthly mean temperatures (2001) from Cavedine meteorological station and related MODIS/Terra data.

Data source	Month of year 2001 1	2	3	4	5	6	7	8	9	10	11	12	Act./Pot. Observ.
Meteo	721	649	722	697	721	697	721	721	697	720	697	721	8484/8760
MODIS	10	32	15	29	23	8	29	35	24	35	32	20	292/730

curves are matching surprisingly well except for January (1), February (2) and July (6). Table 1 reports on the number of observations used for the calculation of the monthly mean temperatures. Only in case of months with nearly continuous cloud cover the mean temperatures deviate significantly. Fig. 2 shows a boxplot which confirms that the data distribution of both data types is similar. A two-sample Kolmogorov-Smirnov test on both distributions results in D=0.1667 and p-value=0.9985, confirming that the distributions significantly match.

It is important to note that LST temperatures are not identical to air temperatures as measured by meteorological stations. However, efforts have been undertaken to derive air temperatures from surface temperatures (e.g., Goetz et al. 2000).

An application related to the distribution of tick-bourne diseases is the calculation of "autumnal cooling". This index is calculated by linear regression to describe the autumnal temperature decline from August to October (northern hemisphere). Sites of tick-bourne encephalitis appear to be characterized by a high rate of autumnal cooling, relative to the annual maximum of the monthly mean LST level in midsummer (Randolph et al., 2000). Using the raster map

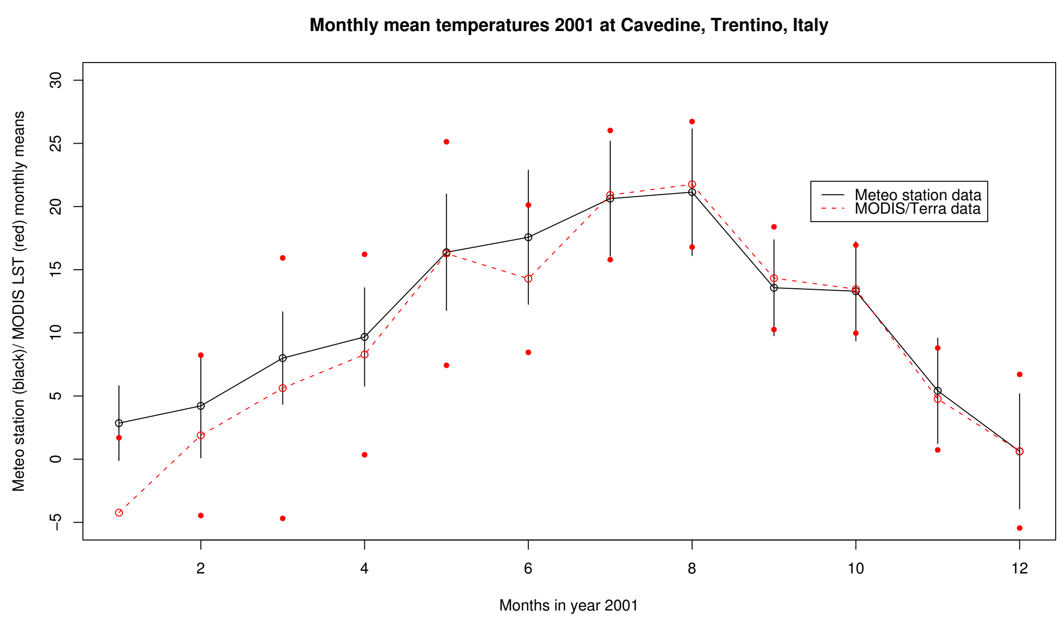

Figure 1: Comparison of monthly mean temperatures (2001) from Cavedine meteorological station (hourly data) and related MODIS/Terra V003 data (at max. two values per day). The standard deviation is indicated (bar for station, dot for MODIS).

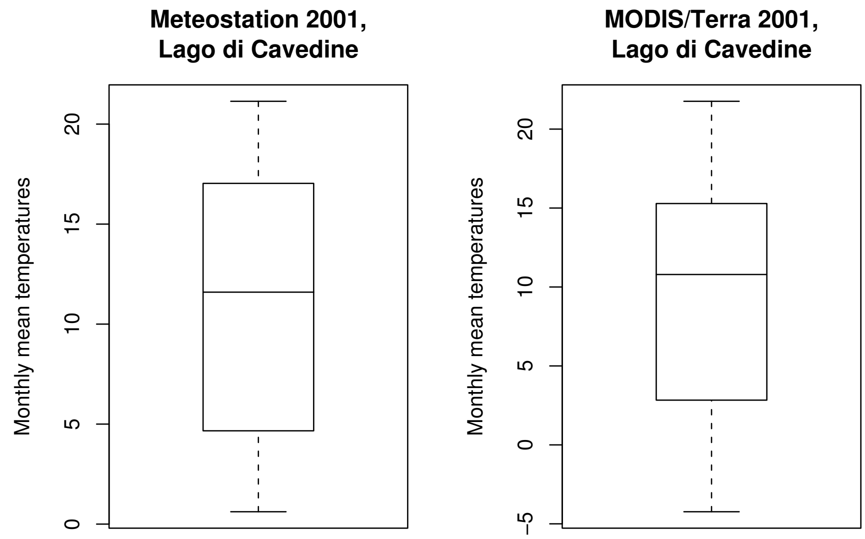

Figure 2: Boxplot of monthly mean temperatures (2001) from Cavedine meteorological station (hourly data) and related MODIS/Terra data (max. two values per day).

time series calculator of GRASS (`r.series`) this index can be easily derived for all years of available MODIS data. This analysis is currently ongoing for the Trentino study subregion.

4.2 Applications of NDVI/EVI 16-day composites

The temporal dynamics of vegetation are an important indicator of vegetation type and, consequently, moisture conditions on the ground. This is an important predictor of suitable habitat for *I. ricinus* (Randolph, 2001). On the one hand NDVI/EVI values can be directly integrated into presence/absence models. On the other hand it is promising to calculate temporal NDVI/EVI differences in order to determine the spatial patterns of spring duration etc. Work is ongoing to relate the density of rodents as tick hosts to the temporal variations of the EVI over the year.

5 CONCLUSIONS AND FUTURE RESEARCH

The usability of MODIS LST and NDVI/EVI data for epidemiological studies appears to be promising. More that 3 years of data are available now, extending the 20 years of time series of AVHRR/NOAA. Data processing to process all available MODIS/Terra data sets in ongoing for the study area. To enhance the temporal resolution of LST maps, we will integrate the MODIS/Aqua data to have four maps per day available (in case of cloud-free conditions).

ACKNOWLEDGEMENTS

M. Neteler was supported by the FUR-PAT Project WebFAQ. The meteorological data have been provided from the Autonomous Province of Trento. Dr. A. Rizzoli/CEA is acknowledged for her support with ticks related data of the ECODIS project.

REFERENCES

Furlanello, C., Neteler, M., Merler, S., Menegon, S., Fontanari, S., Donini, A., Rizzoli, A., and Chemini, C., 2003. GIS and the Random Forest Predictor: Integration in R for Tick-borne Disease Risk Assessment. In Hornik, K. and Leisch, F., editors, *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC2003, Vienna, Austria, March 20-22)*.

Goetz, S., Prince, S., and Small, J., 2000. - In: S.I. Hay and S.E. Randolph and D.J. Rogers, 2000. *Remote sensing and Geographical Information Systems in epidemiology*, chapter Advances in satellite remote sensing of environmental variables for epidemiological applications, pages 289–307. Academic Press.

GSFC/NASA, 2003. MODIS technical specifications. Electronic document: http://modis.gsfc.nasa.gov/about/specs.html.

Hess, G., Randolph, S., Arneberg, P., Chemini, C., Furlanello, C., Harwood, J., Roberts, M., and Swinton, J., 2002. - In: P.J. Hudson and A. Rizzoli and B.T. Grenfell and H. Heesterbeek and A. P. Dobson, 2002. *The Ecology of Wildlife Diseases*, chapter Spatial aspects of disease dynamics, pages 102–118. Oxford University Press.

Huete, A., Justice, C., and Leeuwen, W., 1999. MODIS vegetation index (MOD 13). Algorithm theoretical basis document ATBD13. Electronic document: http://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf.

Merler, S., Chemini, C. F. C., and Nicolini, G., 1996. Classification tree methods for analysis of mesoscale distribution of Ixodex ricinus (Acari: Ixodidae) in Trentino, Italian Alps. *Journal of Medical Entomology*, 33(6):888–893.

Neteler, M. and Mitasova, H., 2004. *Open Source GIS: A GRASS GIS Approach*. Number 773 in SECS. Kluwer Academic Publishers, Boston, 2nd edition.

Randolph, S., 2001. The shifting landscape of tick-borne zoonoses: tick-borne encephalitis and lyme borreliosis in europe. *Phil. Trans. Royal Soc. London B*, 356:1045–1056.

Randolph, S., Green, R., Peacey, M., and Rogers, D., 2000. Seasonal synchrony: the key to tick-borne encephalitis foci identified by satellite data. *Parasitology*, 121:15–23.

Rizzoli, A., Merler, S., Furlanello, C., and Genchi, C., 2002. Geographical information system and bootstrap aggregation (bagging) of tree-based classifiers for lyme disease risk assessment in Trentino, Italian Alps. *Journal of Medical Entomology*, 39(3):485–492.

U.S. Geological Survey, 2004. MODIS Reprojection Tool V3.1. Software: http://edc.usgs.gov/programs/sddm/modisdist/index.shtml.

Wan, Z., 1999. MODIS Land-Surface Temperature. Algorithm theoretical basis document (LST ATBD): LST calculations. Electronic document: http://modis.gsfc.nasa.gov/data/atbd/atbd_mod11.pdf.

Wan, Z., 2003. MODIS Land-Surface Temperature products users' guide. Electronic document: http://www.icess.ucsb.edu/modis/LstUsrGuide/usrguide.html.