

Automated spaceborne detection of degraded vegetation around Monchegorsk, Kola Peninsula, Russia

W. G. Rees

Scott Polar Research Institute, University of Cambridge, Lensfield Road, Cambridge CB2 1ER (wgr2@cam.ac.uk)

ABSTRACT. This paper develops a simple method for the detection of ‘vegetation anomalies’, locations where the amount of vegetation, estimated through the use of the normalised difference vegetation index (NDVI), is significantly lower than expected on the basis of topographic factors alone. The method is developed and tested using satellite imagery from the area around the town of Monchegorsk on the Kola Peninsula, Russia. This area has been subject to heavy levels of airborne industrial pollution for many years and the intended purpose of the method is to allow the extent of pollution damaged vegetation to be estimated with as few operational decisions as possible by the data analyst, thus suiting it for automation and for the analysis of time-series of satellite images. While the work described in this paper is to some extent preliminary, it does establish that spatial variations in the NDVI of undisturbed vegetation can, at least in the study area, be modelled satisfactorily using topographic variables, and that negative departures from this modelled variation are very strongly associated with industrially mediated damage.

Introduction

Industrialisation of the Arctic and sub-Arctic region is far from uniform but, in some locations, rather significant. One particular problem, that has received considerable attention over the last few decades, is the atmospheric emission of sulphur dioxide and heavy metals as a result of the smelting of nickel and other non-ferrous metals. These emissions can produce significant disturbance of vegetation at the local and regional scales. Particular problems are noted in Russia, around the smelters at Monchegorsk and Nikel’-Zapolyarniy on the Kola Peninsula (Rees and Rigina 2003), and especially around Noril’sk in Siberia (Toutoubalina and Rees 1999), and all of these sites have been the subject of considerable study.

Monitoring of pollution impacts by the use of spaceborne remote sensing methods is favourable in these locations (Johansen and others 1995; Rigina 2002; Tømmervik and others 1998). The sensitivity of images acquired in the visible and near-infrared part of the electromagnetic spectrum to the distribution and physiological state of vegetation, especially vascular plants, is well established (Goetz and others 1985), and spaceborne remote sensing offers the possibility of obtaining data from locations that may be difficult to reach for one reason or another, and also of detecting change through the analysis of time series of satellite data that can potentially extend back to the early 1970s. Change-detection methods frequently make use of post-classification analyses (Lunetta and Elvidge 1999), in which each of a series of satellite images is first classified to show the distribution of land cover, including classes representing damaged vegetation, and the classified images are then compared. This approach, however, often encounters a problem in the unavailability of suitable training or validation data from earlier dates. Assumptions about the stability of spectral signatures corresponding to particular land cover types are less likely to hold good at high latitudes than in temperate regions, mainly as a result of

the very rapid phenological progression during the short summer, and the difficulty is often circumvented in other ways, for example by making certain assumptions about the kind of changes that can and cannot take place in the land cover (Rees and Williams 1997). It can be difficult to ensure that such approaches are robust.

The aim of this paper is to develop an approach to the identification of degraded vegetation around a source of atmospheric pollution that is robust in the sense that it requires minimal decision making by the data analyst, and does not require that training data be available. Such a technique would be potentially well suited to investigating the changing distribution of pollution impact over time, using historical satellite data.

Study area

The study area is based around the city of Monchegorsk (67°.9 N, 32°.9 E, 132 metres a.s.l.), in the Murmansk Oblast, Russia (Fig. 1). Monchegorsk (population ca 50,000 at 2002 census) is situated near the northern end of Lake Imandra, in the west-central part of the Kola Peninsula. The area is mountainous, with the Monchetundra ridge to the west and the Khibiny laccolith to the east of Lake Imandra (Fig. 2). These reach elevations of 1000 m and 1300 m respectively. The climate of the Kola Peninsula is determined by atmospheric processes over the Atlantic Ocean. Prevailing winds are from the north in summer, from the south in winter. The mean temperature in the coldest months (January and February) is around –9°C and in the warmest month (August) +12°C; the average annual precipitation is around 500–600 mm (Rees and Kapitsa 1994). Local vegetation is strongly dependent on altitude. The treeline is at around 400 m, below which is boreal forest consisting mainly of fir, spruce and pine. Above the boreal forest is a subalpine zone occupied by birch. The alpine zone consists of dwarf birch (*Betula nana*), lichens, dwarf shrubs and some moss

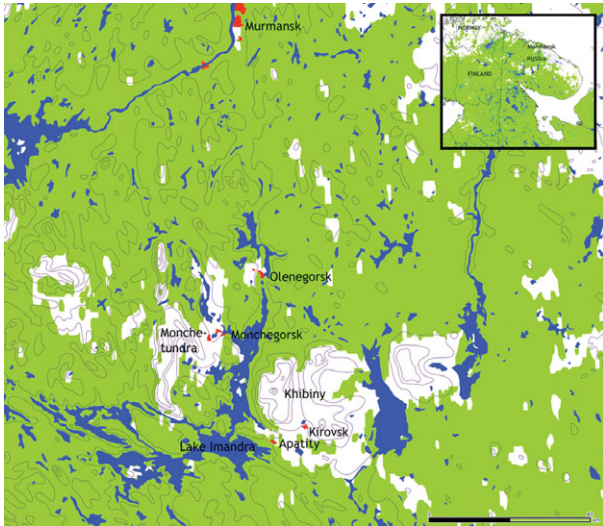


Fig. 1. Location of the study area on the Kola Peninsula, Russia. Green colour denotes forest areas.

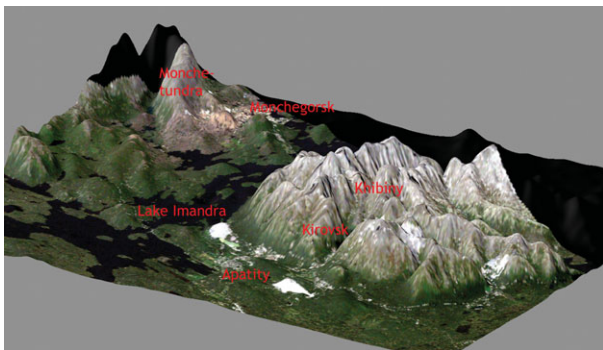


Fig. 2. Landsat-7 Enhanced Thematic Mapper image of the Monchegorsk study area, draped over a Digital Elevation Model. The towns of Monchegorsk, Apatity and Kirovsk are identified, as well as Lake Imandra, the Khibiny mountains, and the mountain ridge Monchetundra. The extensive light areas near Apatity and Kirovsk are industrial tailing ponds. The image was acquired in July 2000.

tundra, while areas above about 800 m are generally devoid of vegetation (Rees and Kapitsa 1994).

Monchegorsk is the site of the Severonikel' nickel smelter, which has been in operation since 1938 and is a major source of acidifying atmospheric emissions and heavy metal pollution (Rees and Rigina 2003). It has been the focus of numerous studies of industrial impact on high latitude vegetation, including by the present author and his colleagues from the Geography Faculty of Moscow State University since 1993.

Fig. 2 shows a satellite image of the study area draped over a Digital Elevation Model of the terrain. The image is a true-colour Landsat-7 image, and it clearly shows the altitudinal zonation of vegetation, especially on the flanks of the Khibiny mountains. In the vicinity of the major source of pollution at Monchegorsk this relationship between elevation and vegetation is perturbed, and this forms the basis of the method developed in this paper.

Proposed method of detecting degraded vegetation

The proposed method is based on the use of the normalised difference vegetation index (NDVI). This is a simple mathematical operation performed on suitable calibrated imagery, and it is essentially a transformed ratio of the scene reflectances r_r and r_i in the red and near-infrared regions of the electromagnetic spectrum (1):

$$N = \frac{r_i - r_r}{r_i + r_r} \quad (1)$$

The NDVI is well known (van Wijk and others 2004) to be strongly correlated to the amount of vascular vegetation (for example the above-ground phytomass) present at a location, as a result of the positive correlation of infrared reflectance and the negative correlation of red reflectance, mediated by cellular structure and the presence of chlorophyll respectively (Tucker and others 1986). It has often been used as a means of identifying vegetation damage (Pitblado and Amiro 1982; Litinsky 1996). The hypothesis tested in the present paper can be stated in two parts: (1) in areas where vegetation is not significantly affected by air pollution, the NDVI of vegetated areas is controlled principally by topographic variables, especially elevation; (2) pollution impact on vegetation is manifested as a lowering of the NDVI relative to the value that would be expected on the basis of the topographic variables. If this hypothesis is substantiated, it offers the promise of a rather simple objective technique for identifying degraded vegetation.

Analysis

The satellite image analysed in this work was an extract of a Landsat-7 Enhanced Thematic Mapper (ETM+) image acquired from path 186 row 013 on 28 July 2000. The image was downloaded as a georeferenced orthoimage (WGS84 datum, UTM zone 36) from the Global Land Cover Facility at the University of Maryland, and trimmed to the region $477888 \leq x \leq 555037.5$, $7482162 \leq y \leq 7541983.5$, where x and y are the UTM coordinates (pixel centres) in metres, that is to an extract 2708 pixels wide and 2100 pixels high (the image pixels were 28.5 m square). The NDVI was calculated from bands three and four of this image using the calibration constants, obtained from the Landsat-7 science data users' handbook (NASA 2011), and the dark-object subtraction method (Vincent 1972) to correct the data for atmospheric effects. This method assumes that the effect of propagation through the atmosphere is simply to add a constant radiance, dependent on the wavelength band, to the detected signal. The appropriate quantity to subtract from the measured radiance is estimated by searching for the pixels contributing least to the at-satellite radiance and assuming that they contribute nothing to the surface-leaving radiance so that what is detected at the satellite is all contributed by the atmosphere. In practice this is a good assumption for pixels corresponding to clear deep water, especially in areas of deep shadow. Such pixels

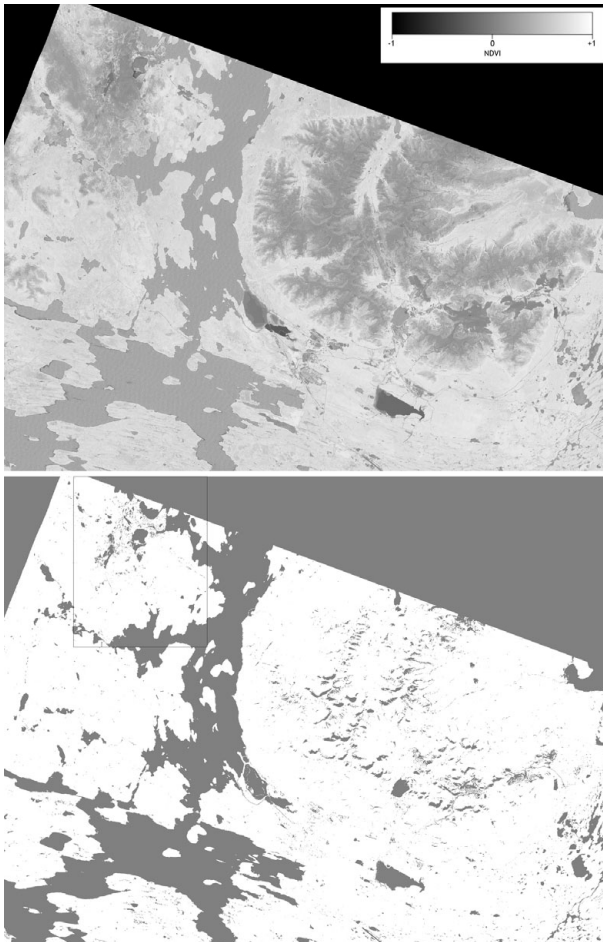


Fig. 3. Above: Calculated NDVI image. Monchegorsk and Monchetundra are towards the top left; the Khibiny mountains are the large, approximately circular feature at right centre. Below: Vegetated areas included in the analysis (white). The 'Monchegorsk box' is shown at the top left.

are not hard to find in this study area. The corresponding NDVI image is shown in Fig. 3.

In order to test the hypothesis, it was necessary first to restrict the analysis only to vegetated, or formerly vegetated, areas. This required the identification of pixels corresponding to water, snow, shadow and tailing ponds, as well as the exclusion of the polluted area around Monchegorsk. Urban areas and roads were not masked. The mask for water, snow etc was generated by performing a 60-cluster unsupervised (ISODATA) classification of the entire image extract and then inspecting each cluster in turn to determine whether it corresponded to one of the classes to be masked. This was straightforward and reasonably objective given the highly distinctive spectra of water, snow and tailings, and only the shadow class presented some ambiguity. This problem was resolved by including any doubtful clusters within the mask, which is shown in Fig. 3. This figure also shows a rectangular region towards the top left of the image extract, centred on Monchegorsk. This has coordinates $486780 \leq x \leq 503880$, $7520095.5 \leq y \leq 7541983.5$

(that is it is 17.1×21.9 km) and will be referred to as the 'Monchegorsk box'. A rectangular region was chosen for simplicity of definition and also having regard to the elongation of the pollution affected region along a north-south axis, presumably as a consequence of the direction of the prevailing wind and also of the constraints on the atmospheric transport of pollutants imposed by the surrounding terrain. The appropriate dimensions of the Monchegorsk box were judged by eye. Data from within the Monchegorsk box were excluded from the determination of the relationship between NDVI and topographic variables for undamaged vegetation. The mean NDVI within the unmasked areas, excluding the data from the Monchegorsk box, was 0.609 with a standard deviation of 0.185.

Topographic variables were derived from Advanced Spaceborne Thermal Emission and Reflection Spectrometer (ASTER) Global Digital Elevation Model (GDEM) data for the study area. ASTER GDEM data are freely available (see Acknowledgements; also Rees (in press)) and are gridded at a sampling interval of one second of arc (approximately 31 m in latitude and 11 m in longitude at the latitude of Monchegorsk), although the effective spatial resolution is around 100 m with an elevation accuracy of around ten m (Rees in press). The data were resampled to the same resolution and projection as the Landsat imagery. Two topographic variables were derived for each pixel in the DEM: the elevation, and the aspect of the slope. Elevation was binned in five m intervals and aspect was binned into the eight cardinal and ordinal directions (north, north-east and so on) relative to grid north.

The NDVI data from the presumed undamaged areas were used to define the variation of the mean NDVI with elevation as a single explanatory variable (Fig. 4). This relationship is as expected in the sense that NDVI decreases almost monotonically with increasing elevation once the elevation exceeds about 300 m. The data plotted in Fig. 4 were not used to define the parameters of a best fitting model of any particular kind, although in fact it was noted that a fourth-order polynomial function gives a reasonable fit.

The root mean square residual of the NDVI data, after subtracting the empirical relationship represented by Fig. 4a, is defined by

$$\langle (N - \bar{N}(h))^2 \rangle^{1/2}, \quad (2)$$

where N is the observed value of the NDVI at a particular location at elevation h , and $\bar{N}(h)$ is the empirical relationship, that is the mean value of N for all undamaged areas with the same elevation. The angle brackets denote averaging over all undamaged areas. This quantity was found to have a value of 0.112. Comparing this to the standard deviation of the NDVI, already noted as 0.185, implies a pseudo- r^2 value of 0.631 ($1 - (0.112/0.185)^2$), that is that around 63% of the variance in the NDVI is explained by this empirical model. Including the effect of aspect as well as elevation increases the pseudo- r^2

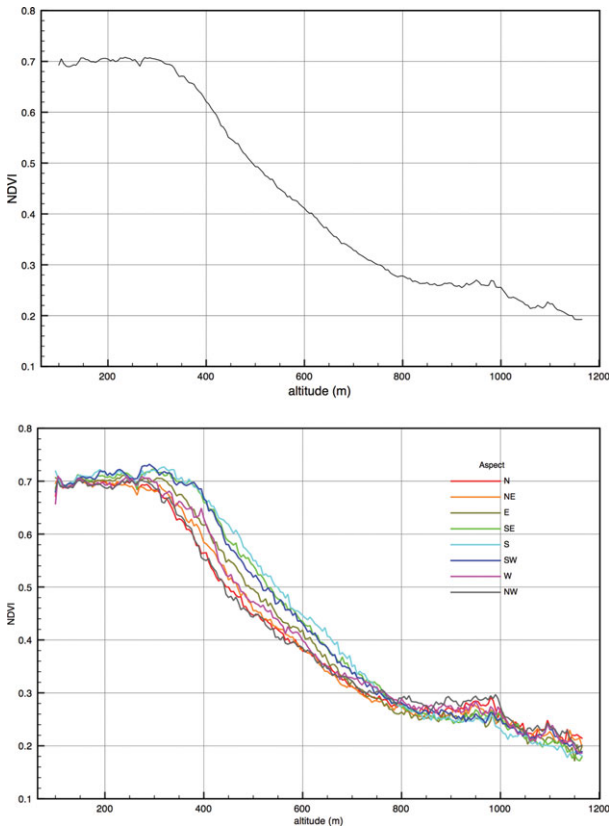


Fig. 4. Dependence of mean NDVI on elevation (above) and on elevation and aspect (below) for presumed undamaged areas.

slightly, to 0.641, and the effect of varying aspect is again as expected, in the sense that south facing slopes have higher values of mean NDVI than north-facing slopes (Fig. 4). However, the differences are not large compared with the rms residual, so the effect of aspect was ignored in subsequent analysis. The NDVI anomaly was thus defined as

$$\Delta N = N - \bar{N}(h). \tag{3}$$

An NDVI anomaly image was created for all unmasked areas, including the Monchegorsk box.

Results

The NDVI anomaly image resulting from this procedure is shown in Fig. 5. Where there are significant anomalies ($|\Delta N| > 0.2$) these are almost always negative. Strong negative anomalies ($\Delta N < -0.4$) are, without exception, associated with urban areas, quarries, tailings or the region close to the smelter at Monchegorsk, and it seems very likely that these strong negative anomalies can be identified as areas of significantly degraded vegetation. Negative anomalies as strong as -0.6 are observed close to Monchegorsk. The weaker negative anomalies also appear to be associated with the degradation of vegetation around Monchegorsk, but they are also less believably associated with the lower slopes of the Khibiny mountains.

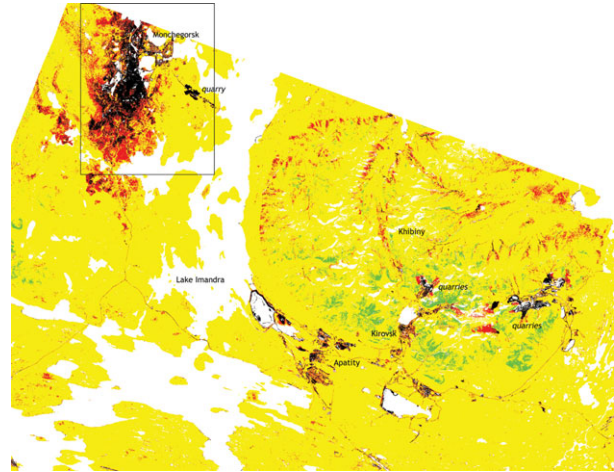


Fig. 5. NDVI anomaly relative to the elevation-dependent model of figure 4. Black: $\Delta N < -0.4$; red: $-0.4 \leq \Delta N < -0.2$; yellow: $-0.2 \leq \Delta N < +0.2$; green: $\Delta N \geq +0.2$. The location of the Monchegorsk box is shown.

Some positive anomalies occur in the southern part of the Khibiny mountains.

Discussion

The method developed in this paper appears to have successfully identified the distribution of degraded vegetation around the smelter at Monchegorsk on the basis that the value of the NDVI in the damaged area are significantly lower than would be expected in the absence of disturbance. The method relies on the fact that the value of the undisturbed NDVI is strongly controlled by the elevation, presumably reflecting the fact that the study site is close to the treeline. Confirmation that it is reasonable to associate negative NDVI anomalies with degraded vegetation was provided by some *in situ* verification. Twenty locations were visited during field-work between 2001 and 2009. At each location, site descriptions were completed and photographs were taken. Fig. 6 shows photographs from four of these locations, chosen to illustrate two locations at similar elevation (and hence having similar values of \bar{N}) but with substantially different values of N , and two locations having similar values of N at substantially different elevations.

The top two photographs in Fig. 6 show two sites that would be expected to have similar NDVI values. The first is from a location about 20 km south of Monchegorsk, with largely undamaged pine and spruce forest, while the second is from a location about 10 km south of Monchegorsk showing heavily damaged forest with a small amount of birch scrub. The NDVI anomalies at these two locations are $+0.08$ and -0.52 respectively. The bottom two photographs show two sites with very similar values of the observed NDVI. The first of these is at an elevation of 547 m and is a healthy mixed tundra of dwarf shrubs and lichens in the Khibiny mountains, while the second is of moderately damaged forest about 30 km south of Monchegorsk.



Fig. 6. Field photographs of sites having the following characteristics, in each case presented as elevation in metres, followed by modelled NDVI \bar{N} and observed NDVI \bar{N} . Top left: 207, 0.70, 0.78; top right: 205, 0.70, 0.18; bottom left: 547, 0.44, 0.48; bottom right: 160, 0.71, 0.47.

The methodology developed in this paper has used calibrated satellite imagery to calculate the NDVI. However, it would be possible to apply a variant of the method even if the calibration data were not known. Specifically, we suppose that a ‘pseudo-NDVI’ N' is calculated from uncalibrated reflectance or radiance data as follows:

$$N' = \frac{r'_i - r'_r}{r'_i + r'_r} \quad (4)$$

where r'_i and r'_r are monotonic functions of the calibrated reflectances r_i and r_r respectively, and $r'_i = 0$ when $r_i = 0$ and $r'_r = 0$ when $r_r = 0$. In this case, which requires only that dark-object subtraction is carried out correctly on the uncalibrated data, it can be shown that N' is related monotonically, although not linearly, to N . Thus, provided a monotonic relationship is found between the mean pseudo-NDVI N' and the elevation, anomalous values of N' can be interpreted as though the vegetation cover is unexpected for the elevation at which they occur, and they can be expressed as elevation anomalies. This principle could of course equally well be applied to calibrated data. For example, the field location shown at the top right of figure 6 has an actual value of NDVI (0.18) that is, according to Fig. 4, expected for an elevation of above 1165 m, while the actual elevation of the location is 205 m, so the elevation anomaly is greater than +945 m.

The field location shown at the bottom left of Fig. 6 has an actual value of NDVI (0.48) which corresponds to an elevation of 515 m. Since the actual elevation is 547 m, the elevation anomaly is –32 m.

The use of the ‘Monchegorsk box’ to eliminate most of the area of degraded vegetation from the characterisation of the relationship between elevation and NDVI relied on prior knowledge of the study area. However, this step could be omitted if the data met certain conditions. What is needed is to be able to identify the NDVI corresponding to undamaged vegetation at each elevation, in the presence of NDVI values corresponding to degraded vegetation. Provided that undamaged vegetation is substantially more common than damaged vegetation at every elevation considered in the analysis, the problem can be addressed by statistical means. One obvious possibility would be to estimate the modal value of the NDVI in each range of elevation; another would be to fit a suitably chosen function with parameters representing both undamaged and degraded vegetation.

There is a relevant sense in which the vegetation in the study area is rather uniform. Although an increase in elevation from 300 to 1200 m sees a series of transitions in the baseline vegetation, from boreal forest, to forest tundra, to dwarf-shrub-lichen tundra and ultimately to predominantly bare ground, Fig. 4 strongly suggests that

these transitions can all be regarded as part of a continuous decrease in the amount of vascular vegetation. This is probably in part a consequence of the relatively coarse spatial resolution of the satellite imagery, so that all pixels are mixed and most contain at least some vascular plant material, but it is perhaps also a consequence of the relative simplicity of the ecosystems. Phenological considerations may also be important: the image analysed here may by chance have been acquired during the optimum point in the phenological cycle, with maximum greenness of the vegetation enhancing the elevational differences and those attributable to industrial impact (Tutubalina and Rees 2001). These points could be tested by analysing images from other test sites and also by analysing a time-series of images from the same site: these analyses are currently being undertaken.

Conclusions

The data analysed in this paper strongly suggest that (a) the spatial variation in the NDVI of relatively undisturbed vegetation can be explained rather satisfactorily using elevation as a single explanatory variable, and (b) negative departures from this relationship can be used to identify areas in which vegetation is notably degraded. The approach developed here has considerable potential for investigating the development of the region of damaged vegetation around the nickel smelter at Monchegorsk over the past few decades, and may also be extensible to other sites.

Acknowledgements

I am grateful to my many friends and colleagues at Moscow State University with whom I have conducted fieldwork around Monchegorsk since 1993, and especially to Olga Tutubalina. ASTER GDEM data are a product of METI and NASA. Data were downloaded from Earth Observing System Data and Information System (EOS-DIS). 2009. Earth Observing System Clearing House (ECHO)/Warehouse Inventory Search Tool (WIST) Version 10.X [online application]. Greenbelt, MD: EOS-DIS, Goddard Space Flight Center (GSFC) National Aeronautics and Space Administration (NASA). URL: <https://wist.echo.nasa.gov/api/>. Landsat satellite imagery was downloaded from the Global Land Cover Facility, University of Maryland. URL: <http://www.glcfc.umd.edu>. Data processing was conducted on an Apple Macintosh computer using free software: ImageJ (Rasband 1997–2011) and Multispec (Biehl and Landgrebe 2002).

References

Biehl, L., and D. Landgrebe. 2002. MultiSpec – a tool for multispectral-hyperspectral image data analysis. *Computers and Geosciences* 28 (10): 1153–1159.

- Goetz, A.F.H., G. Vane, J.E. Solomon, and B.N. Rock. 1985. Imaging spectrometry for Earth remote sensing. *Science* 228: 1147–1153.
- Johansen, M.E., H. Tømmervik, T. Guneriusson, and J.P. Pedersen. 1995. Using a geographic information system (ARC/INFO) to integrate remote sensed and *in-situ* data in an analysis of the air pollution effects on terrestrial ecosystems in the border areas between Norway and Russia. In: Lobersli, E., and K. Venn (editors). Second symposium on the effects of air pollutants on terrestrial ecosystems in the border area between Russia and Norway. Svanvik, Norway: Statens forurensningstilsyn: 120–130.
- Litinsky, P. (editor). 1996. Assessment of boreal pine forests decline around ore/dressing mill using satellite images. Petrozavodsk: Russian Academy of Sciences (in Russian).
- Lunetta, R., and C. Elvidge. 1999. *Remote sensing change detection*. London: Taylor & Francis.
- NASA (2011). URL: <http://landsathandbook.gsfc.nasa.gov/handbook>. (accessed 24 May 2011).
- Pitblado, J.R., and B.D. Amiro. 1982. Landsat mapping of the industrially disturbed vegetation communities of Sudbury, Canada. *Canadian Journal of Remote Sensing* 8(1): 16–29.
- Rasband, W.S. 1997–2011. ImageJ. US National Institutes of Health, Bethesda, Maryland, USA. URL: <http://imagej.nih.gov/ij> Last (accessed 24 May 11).
- Rees, W.G., and A.P. Kapitsa. 1994. Industrial pollution in the Kol'skiy Poluostrov, Russia. *Polar Record* 30: 181–188.
- Rees, W.G., and O. Rigina. 2003. Methodologies for remote sensing of the environmental impacts of industrial activity in the Arctic and Sub-Arctic. In: Rasmussen, R.O., and N.E. Koroleva (editors). *Social and environmental impacts in the north*. Dordrecht: Kluwer: 67–88.
- Rees, W.G., and M. Williams. 1997. Monitoring changes in land cover induced by atmospheric pollution in the Kola Peninsula, Russia, using LANDSAT MSS data. *International Journal of Remote Sensing* 18: 1703–1723.
- Rigina, O. 2002. Environmental impact assessment of the mining and concentration activities on the Kola Peninsula, Russia, by multirate remote sensing. *Environmental Monitoring and Assessment* 75(1): 13–33.
- Toutoubalina, O.V., and W.G. Rees. 1999. Remote sensing of industrial impact on Arctic vegetation around Noril'sk, northern Siberia: preliminary results. *International Journal of Remote Sensing* 20: 2979–2990.
- Tømmervik, H., M.E. Johansen, J.P. Pedersen, and T. Guneriusson. 1998. Integration of remote sensed and *in-situ* data in an analysis of the air pollution effects on terrestrial ecosystems in the border areas between Norway and Russia. *Environmental Monitoring and Assessment* 49: 51–85.
- Tucker, C.J., I.Y. Fung, C.D. Keeling, and R.H. Gammon. 1986. Relationship between atmospheric CO₂ variations and a satellite-derived vegetation index. *Nature* 319: 195–199.
- Tutubalina, O.V., and W.G. Rees. 2001. Vegetation degradation in a permafrost region as seen from space: Noril'sk (1961–1999). *Cold Regions Science and Technology* 32(2): 191–203.
- Vincent, R.K. 1972. An ERTS multispectral scanner experiment for mapping iron compounds. In: 8th international symposium on remote sensing of environment. Ann Arbor, Michigan: Environmental Research Institute of Michigan: 1239–1247.
- van Wijk, M.T., M. Williams, and G.R. Shaver. 2004. Tight coupling between leaf area index and foliage N content in arctic plant communities. *Ecosystem Ecology* 142 (3): 421–427.