Portability of neural nets modelling regional winter crop acreages using AVHRR time series

Clement Atzberger1* and Felix Rembold2

1Institute for Surveying, Remote Sensing and Land Information (IVFL), University of Natural Resources and Life Sciences (BOKU), Peter Jordan Str. 82, 1190 Vienna, Austria
2Joint Research Center of the European Commission, Via Enrico Fermi 2749, 21027 Ispra, Italy
*Corresponding author, e-mail address: clement.atzberger@boku.ac.at

Abstract
Time series of coarse resolution imagery offer the advantage of free global coverage but have to deal with mixed pixels. The study uses neural nets as modelling tool for sub-pixel crop acreage estimation. Nets are trained with reference crop acreage information derived from 30 m Landsat images and CORINE LC map for interpreting changes in the shapes of coarse resolution AVHRR NDVI profiles. Using official AGRIT statistics for Tuscany (Italy) as reference information, the network portability across years was evaluated. Using 3 images acquired before 2002 nets were trained. Subsequent application to 2002-2009 data explained roughly half of the inter-annual variance.

Keywords: Soft classification, non-linear unmixing, portability, NOAA-AVHRR, neural network, crop acreage.

Introduction
Consistent and timely information about crop acreages are required for assessing the agricultural production of a given area and associated planning purposes. Such information can in principle be generated by supervised classification of high resolution images acquired at key phenological stages. These approaches, however, are labour and cost intensive, and require amounts of cloud-free high spatial resolution imagery, prohibitive for operational implementation over large areas and in multiple years [Lobell and Asner, 2004]. Data availability – in particular if specific crop stages need to be imaged – is often insufficient [Annoni and Perdigao, 1997]. Additionally, these methods are usually affected by misclassification errors which can lead to strong bias in area estimates [Czaplevsky, 1992; Gallego, 2004]. These factors have limited the possibility of automatically updating land cover maps over large areas at regular (annual) intervals [Chang et al., 2007]. For example, for the US it is expected that the Landsat-based NLCD land cover data base will result in approximately 6-year delay between data collection and product availability.
Low resolution images, on the other hand, are easily processed on a regional scale, provide consistent information at high temporal frequency back to the 1980’s and cover large areas at low costs. Therefore, maps derived from such data can be updated at annual intervals [Lunetta et al., 2010]. However, because of subpixel heterogeneity, the application of traditional hard classification approaches cannot be recommended and may result in significant errors in the estimated crop areas [DeFries et al., 1995; Chang et al., 2007]. Significant improvements for crop area estimation at the regional level have been obtained by the introduction of the MODIS sensor with 250 m ground resolution as shown by the results of Lunetta et al. [2006, 2010], Chang et al. [2007], Wardlow and Egbert [2008] and Fritz et al. [2008] amongst others. Not surprisingly, the best results of these studies were obtained for agricultural areas such as the central plains of the United States and the Don river basin in Russia, where typical field sizes are large enough to be resolved by the 250 m spatial resolution. To overcome the problem of subpixel heterogeneity common for many areas of the world with fragmented landscapes, Quarmby et al. [1992] employed linear mixture model (LMM) techniques to coarse resolution data. Hansen et al. [2002] employed the continuous field algorithm for mapping vegetative traits such as tree cover using MODIS data. In the continuous field approach, each coarse resolution pixel is characterized as 0 to 100 percent cover of a vegetation class, ameliorating the primary limitation of coarse spatial resolution data [Chang et al., 2007].

A probabilistic linear unmixing approach with MODIS spectral/temporal data was developed and tested by Lobell and Asner [2004], termed probabilistic temporal unmixing (PTU). The approach estimates subpixel fractions of crop area based on the temporal reflectance signatures throughout the growing season. Endmember sets representing the full range of potential variability are constructed using Landsat data so as to identify pure pixels, mainly located within large fields. The uncertainty in endmember fractions arising from endmember variability can then be quantified using Monte Carlo techniques. The performance of the proposed approach was assessed over Mexico and the Southern Great Plains and varied depending on the scale of comparison. Coefficients of determination ranged from greater 80% for crop cover within areas over 10 km² to roughly 50% for estimating crop area within individual MODIS pixels.

Several authors combined higher resolution images with NOAA AVHRR 1-km imagery to improve sub-pixel crop monitoring capabilities [Maselli et al. 1998; Doraissamy et al. 2004]. However, insufficient contrast between endmembers leads to unstable solutions, resulting in inaccurate fraction images [Lobell and Asner, 2004]. On the other hand, too few endmembers will fail to correctly represent the input signature.

Several studies used Spectral Angular Mapping (SAM) for measuring inter-annual crop area changes based on NDVI time series from NOAA-AVHRR [Rembold and Maselli, 2004, 2006]. The studies found that it was feasible to derive relatively accurate inter-annual winter crop acreage changes for the region of Tuscany, Italy. However, good results could be obtained only by estimating the crop acreage changes of single years from the average of a high number (seven) of reference years, while the results were significantly worse by using less or single years of reference data. No attempts were undertaken to take benefit of available information about the fractional coverages of non-arable land end-members.
Lunetta et al. [2010] used MODIS 16-day composite data to develop annual cropland and crop-specific map products for the Laurentian Great Lakes Basin. For the categorization of cropland and crop types a multi-layer neural network classifier was used. Training pixels were identified using visual interpretation of MODIS-NDVI temporal profiles. The research included the application of an automated protocol to first filter the NDVI data to remove poor (corrupted) data values and then estimate the missing data using a discrete Fourier transformation technique. The use of neural nets for estimating sub-pixel land cover from temporal signatures was investigated by Karkee et al. [2009]. Braswell et al. [2003] demonstrated that network based non-linear regression offers significant improvement relative to linear unmixing for estimation of sub-pixel land cover fractions in the heterogeneous disturbed areas of Brazilian Amazonia. The improvement was related to the fact that linear unmixing assumes the existence of pure sub-pixel classes (end-members) with fixed reflectance signatures. The neural network approach proposed by Braswell et al. [2003] estimates nonlinear relationships between each land cover fraction and spectral-directional reflectances, without making assumptions about the physics of sub-pixel mixing. In Verbeiren et al. [2008] monthly MVC of SPOT-VGT (between March and October) were used to model the area fraction images (AFI) of eight classes in 2003 for Belgium. Relatively good results were obtained especially if the initial (pixel-based) results were aggregated to higher regional levels. The portability across growing seasons was investigated by Bossyns et al. [2007] in an accompanying paper on the same data set. The nets were trained on data of one growing season and then applied to SPOT-VGT of the training year plus three additional seasons. High and stable accuracies of the estimated AFI’s were obtained for the training data. For example, at regional level, the $R^2$ for winter wheat of the training years was ~0.8 (0.67-0.87). On average, however, this values decreased by 0.45 units when the networks were applied to different seasons, probably because of a (too) high inter-annual variability of the ‘end-members’.

To better cope with the natural year-to-year variability of NDVI profiles of vegetated surfaces, Atzberger and Rembold [2009, 2010] trained networks (in a leave-one-out mode) with AVHRR imagery acquired over eight years. The target variable represented the sub-pixel winter crop fractional coverage. To permit the net distinguishing for various proportions of non-arable land within the mixed pixels (e.g. forested areas, urban land, etc.), available land cover information was used as additional input. The research demonstrated that a considerable amount of the spatial variability of winter crop surfaces (cross-validated $R^2$ ~0.7-0.8) can be mapped using time series of coarse resolution imagery. A positive impact was demonstrated regarding the concurrent use of ancillary information and the applied temporal smoothing algorithm. In-season predictions improved compared to the work of Rembold and Maselli [2004, 2006] using the same data set and linear prediction models.

The current paper is an extension and refinement of the mentioned conference papers [Atzberger and Rembold, 2009, 2010]. Its main objective is to test if neural networks can be fed with low resolution AVHRR imagery and additional ancillary information (derived from CORINE land cover) to estimate reliably, and across growing seasons, regional crop acreages of winter wheat. For evaluating the portability of networks across growing seasons - in a realistic/feasible setting - nets were trained with three high resolution images acquired between 1988 and 2001. Once trained, the nets were applied to the following time period (2002 to 2009) for estimating the regional winter crop acreage of Tuscany. The estimates were validated against official agricultural statistics.
Material
The methodology was tested over the Tuscany region in Central Italy. The choice was driven mainly by the availability of both satellite imagery and agricultural statistics. The region is covered by a consistent NOAA-AVHRR data time series taken in the period from 1986 to 2009, when also several Landsat TM/ETM+ scenes were acquired. An area frame sampling method [Carfagna et al., 1998] has been regularly applied since 1988 to measure the extent of the main crops in Tuscany.

Figure 1 - (left) Example AVHRR-NDVI image of the study area Tuscany. The inlet shows the location of Tuscany within Italy. (center) Average climatic conditions for Tuscany’s capital city Florence. (right) Generalized dominant slope class map of Tuscany (source: European Soil Database).

Geography and environmental features of the study area
Tuscany is situated between 9°-12° East longitude and 44°-42° North latitude, covering circa 2 x 10^6 hectares (Fig. 1). From an environmental point of view, Tuscany is peculiar for its extremely heterogeneous morphological and climatic features. The topography ranges from flat areas near the coast-line and along the principal river valleys to hilly and mountainous zones towards the Apennine chain. Approximately 2/3 of the region is covered by hilly areas, 1/5 by mountains and only 1/10 by plains and valleys (Fig. 1 right). From a climatic viewpoint, Tuscany is influenced by its complex orographic structure and by the direction of the prevalent air flows (from West / North-West). As a result, the climate ranges from typically Mediterranean to temperate warm or cool according to the altitudinal and latitudinal gradients and the distance from the sea. The average climatic conditions for the city of Florence are depicted in Figure 1 (centre). The land use of Tuscany is predominantly agricultural where the land is relatively flat (slope < 15%) (Fig. 2) and where soil organic carbon content is high. In hilly and mountainous areas mixed agricultural and forestry dominate. The main agricultural cover types are cereal crops in the plains and olive groves and vineyards on the hills. The upper mountain zones are almost completely covered by pastures and forests. Cropland is spread over the coastal zones, the inner plain and gently sloping hilly
areas and covers approximately 20% of the land surface (Fig. 2). The prevalent cereal is durum wheat, with an average planted area of 110,000 ha and with a mean growing cycle from November to the end of June (Fig. 3). In areas with low percentage of arable land, forests and plantations occupy the main remaining area. When arable land dominates, the remaining area is mainly occupied by grass- and shrubland, respectively urban land. This land cover and land use association is further detailed in Table 1.

Table 1 - Average land occupation (in percent) within five arable land (AFIw) classes (fractional coverages of 0-20, 20-40, …, 80-100%). The values indicate the number of pixels in relation to the total number of pixels within the AFIw class (in percent). The relative number of AVHRR pixels for each of the five arable land AFIw classes is indicated in the first column (in parentheses) in relation to the total pixel number of the study region.

<table>
<thead>
<tr>
<th>AFIw class arable land (in %)</th>
<th>forest</th>
<th>plantation</th>
<th>grass &amp; shrubland</th>
<th>urban &amp; bare soil</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;0-20 (11.7%)</td>
<td>34.9</td>
<td>7.6</td>
<td>39.8</td>
<td>7.4</td>
</tr>
<tr>
<td>20-40 (7.9 %)</td>
<td>23.7</td>
<td>5.3</td>
<td>33.1</td>
<td>6.3</td>
</tr>
<tr>
<td>40-60 (7.2%)</td>
<td>15.8</td>
<td>3.1</td>
<td>24.6</td>
<td>5.6</td>
</tr>
<tr>
<td>60-80 (6.7%)</td>
<td>8.2</td>
<td>1.4</td>
<td>14.7</td>
<td>5.2</td>
</tr>
<tr>
<td>&gt;80 (11.1%)</td>
<td>1.3</td>
<td>0.3</td>
<td>3.0</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Figure 2 - Maps displaying the spatial distribution of area fraction images (AFI) of major CORINE land cover classes resampled to AVHRR pixel size: forest, arable land, grass- and shrubland, plantation and urban/bare land. The two pie charts in the central bottom part of the figure indicate the general land cover of Tuscany (large) and the distribution of five AFIw classes within the arable land (small). All data based on CORINE.
Figure 3 - NDVI profiles of winter crops (in black) and the remaining crop land (labelled “summer crops”) (in red) in Tuscany derived from pixels with a high proportion of arable land (CORINE). The white areas highlight the dekads used for modelling (7-13 and 17-25).

Data

Reference information on winter crop areas

Wheat area estimates of Tuscany for the period 1988-2009 were obtained from the AGRIT project [Consorzio ITA, 1987; AGRIT, 2009]. These statistics are produced annually through an area frame sampling method, which guarantees high estimation accuracy at the regional scale. At the regional level, the error for the main crops is generally lower than 10% [Carfagna et al., 1998]. In what follows, we use the term “winter crop area” as a synonym for the wheat area. The term “summer crop” designates the remaining crop land within the CORINE class “arable land”. Figure 4 displays the AGRIT winter crop area for the region of Tuscany between 1988 and 2009. A considerable variability can be seen, mainly as a direct result of European agricultural policies, market prices and weather conditions. The land cover classification of Tuscany produced by the CORINE project was used (i) for deriving high resolution reference maps showing the spatial distribution of winter crops and (ii) as secondary input for the neural net. The Tuscan Regional Service for Cartography provided the necessary data. Data came in the form of a digitized map with a nominal scale of 1:100 000. The methodologies used for the production of the map as well as its main features are described by Annoni and Perdigao [1997].

High resolution images

High resolution images were used for spatializing and disaggregating the statistical information provided by the AGRIT statistics. The high resolution data set consisted of 8 Landsat frames (192/30), 6 taken by the Thematic Mapper (TM) (1988, 1991, 1992, 1995, 1997 and 1998) and 2 by the Enhanced Thematic Mapper (ETM+) (2000, 2001). All of them were acquired during the month of August and were cloud free over the main agricultural areas. The Landsat TM/ETM+ scenes were first georeferenced by a nearest neighbor resampling
algorithm using about 120 control points selected on a land/water mask derived from the CORINE map. TM Bands 4 (nIR) and 3 (Red) were corrected for atmospheric effects and converted into reflectance images following the method proposed by Gilabert et al. [1994]. The reflectances of these two bands were used to compute a high resolution NDVI image for every training year. The years for which Landsat TM/ETM+ imagery was available are highlighted in Figure 4 (designated as “data set 1” – large squares). Note that high resolution reference information does not cover the full range of possible winter crop acreages.

Low resolution images
The Joint Research Centre in Ispra, Italy (JRC-MARS) of the European Commission owns the most elaborate archive of NOAA-AVHRR 1km data over the pan-European continent. The raw data were first collected by JRC itself, and since 2001 by VITO (Belgium). In 2008, all historical data were re-processed with new procedures, which resulted in an unique archive of 29 years [Weiss et al., 2010]. For this study, the AVHRR time series from 1988 to 2009 was used. At the time where the analysis was conducted, the recently updated NOAA AVHRR calibration coefficients [Molling et al., 2010] were not yet available.

Methods
The flowchart in Figure 5 highlights the overall methodology of the study. Main processing steps relate to the data preparation (time series filtering and generation of reference area fraction images) and will be described first, followed by a description of the main modelling task (neural networking).
Figure 5 – Methodological outline: (top left) filtering of the 22-year AVHRR time series. (top right) generation of area fraction images (AFI) using AGRIT statistics, Landsat TM/ETM+ imagery (8 years) and CORINE land cover (CLC) map. First, 5 primary cover type classes are derived from CLC: arable land (A), forest (F), plantation (P), grass- and shrubland (G) and urban/bare land (U). The maps are spatially degraded (AFI calc) to yield AFI at 1 km resolution (AFIa, AFIf, ..., AFIu). In a second step, 8 winter crop masks are derived. This is achieved through optimization of a NDVI threshold applied to the Landsat imagery masked by the arable land mask (A) and using the corresponding AGRIT reference information. Thus, for each of the 8 years, a high resolution mask of winter crops is obtained, next spatially degraded to 1 km AFIw images. (bottom) network training and validation using AVHRR time series and 5 AFI images as inputs and predicting the winter crop area fractions (AFIw). For training only reference information from 3 out of 8 years is used. After training, the net is fed with the AVHRR data (2002-2009) to provide estimates winter crop acreage. Results are checked against the available (AGRIT) reference information (STAT).
Smoothing of low resolution AVHRR data

Time series from AVHRR and other low to medium resolution sensors require a careful filtering/smoothing before they can be applied within quantitative studies [Lunetta et al., 2006]. The standard maximum value compositing (MVC) only corrects for major disturbances. To filter such data, a number of smoothing techniques are available [e.g., Chen et al., 2004; Beck et al., 2006; Atkinson et al., 2012]. For the current study, the Whittaker filter was chosen [Eilers, 2003]. The Whittaker smoother is very fast and automatically optimizes its smoothness parameter using a built-in jackknife procedure. It also handles easily (large stretches of) missing values. In a recent comparative analysis [Atkinson et al., 2012] the filter compared well with competing approaches, including Fast Fourier Transform (FFT), double logistics and asymmetric Gaussian functions.

To deal with negatively biased noise typical for NDVI time series, the original Whittaker smoother was slightly modified [Atzberger and Eilers, 2011]. Negatively biased noise results for example from undetected (sub-pixel) clouds and poor atmospheric conditions, depressing the observed NDVI values. Sudden drops in NDVI must be regarded as noise and removed as they are not compatible with the gradual process of vegetation growth and decline. The modification of the Whittaker smoother allows fitting the upper envelope of the data. Such an iterative fitting was suggested by Beck et al. [2006] amongst others. In the chosen implementation, the smoother first filters the time series using the basic algorithm. Next, all observations that lie below the fitted curve (plus those that were missing and/or quality flagged) are set to the value of the smoothed curve at the given time stamp. With the updated values, the smoothing is repeated. For the third (and final) iteration, the values below the curve are again replaced by the smoothed values followed by a final run of the filter.

For providing crop acreage information not later than September-October of the current year, from the filtered AVHRR data dekads 7 to 13 and 17 to 25 were extracted as inputs to the neural net (highlighted in Fig. 3). During dekads 14 to 16, winter and “summer” crop signatures strongly overlap. These dekads were therefore excluded from the modeling. Only this reduced and smoothed subset was used in the study.

Generation of reference abundance maps using Landsat imagery

The proposed network uses AVHRR NDVI time series and general land cover information (e.g., primary cover types) as input. For network training, one has also to provide reference abundance maps of winter crop acreages (AFIw) as target variable (Fig. 7). The necessary information was obtained by grouping the 42 CORINE land cover classes into environmentally meaningful categories with more or less similar NDVI profiles [Annoni and Perdigao, 1997]. In addition to the “arable land” class, four primary cover types were derived from the CORINE map as “forests”, “pastures”, “tree plantations” and “urban areas” [Maselli, 2001] (Fig. 2). For the study, we assumed that the five primary classes remain stable over the time period covered by the low resolution data set (1988-2009) and that the sum of these five classes sum up to unity. The fractional coverage of winter crops within the coarse resolution AVHRR grid was generated using available TM/ETM+ images of the eight training years (1988, 1991, 1992, 1995, 1997, 1998, 2000 and 2001) and splitting/disaggregating the CORINE “arable land” class into “winter crops” and remaining “summer” crops. To achieve this splitting, a simple NDVI thresholding approach was applied to each individual TM/ETM+ image. The rationale
for this operation was that winter crop fields are almost bare in August, while other fields with summer crops (maize, sunflower, etc.) but also fallows and pastoral areas are in a “green” phase (Fig. 3). Hence, any bare soil in August belongs to the class “winter crop”. The class “summer crop”, on the other hand is a mixture of typical summer crops with other classes. As outlined, a NDVI threshold was optimized for each high resolution TM/ETM+ image to find the most appropriate cutting point distinguishing between bare soils and vegetation. To derive the spatial distribution of winter crops within the CORINE class “arable land”, different NDVI thresholds were systematically evaluated (0.1 ≤ threshold ≤ 0.25). For each threshold the resulting (total) winter crop acreage was compared with the official AGRIT statistics. The threshold minimizing the difference between the reference information and the calculated winter crop acreage was retained and used to derive the high resolution winter crop masks. This was done individually for each of the eight available high resolution images.

The six categories identified in the eight training dates (5 CORINE categories plus “winter crops”) were next spatially degraded by pixel averaging to produce abundance images (i.e., area fraction images) with the same spatial resolution as the AVHRR images (Fig. 2). Out of these maps only the winter crop abundance maps (AFIw) were different for each year, as the other (CORINE) categories were assumed stable over time. The inter-annual variability of the reference winter crop acreages is shown in Figure 6 for all AVHRR pixels with at least 2 percent arable land.

It has to be noted that the TM/ETM+ frame 192/30 does not cover the entire region of Tuscany. Some small areas in the northern part of the region are not covered by the frame.
These areas were therefore removed from the analysis, which corresponded to masking out 3.3% of the “arable land” class. The resulting error was regarded insignificant for the current study.

**Neural networking**

Neural networks (NN) are powerful modelling tools, as they can learn and represent any kind of (non-linear) relationship between (a set of) input variables and (one or several) output variables provided enough nodes are used in the so-called hidden layer [Atkinson et al., 1997; Atkinson and Tatnall, 1997]. Network training consists in confronting a randomly initialized neural net of a given structure with training samples (= inputs and targets). In an iterative way, the values of weights and biases in the network are adjusted so to minimizing the error between the network output and the presented reference targets (here: winter crop abundances at pixel level). Once trained, even very large input data sets can be processed efficiently to produce the requested outputs [Mas and Flores, 2008; Foody et al., 1997].

For the present study, a simple NN with one hidden and one output layer was used (Fig. 7). The output layer represented the winter crop fraction within a coarse resolution pixel and had thus only one node. Twenty-one input nodes were selected for modelling: 5 (inter-annually constant) abundance values derived from the CORINE data (e.g., arable land, forest, tree plantations, pasture and urban) (Fig. 2), plus 16 nodes covering the NDVI profiles of the winter growing season (from dekad 7 to dekad 25), excluding the dekads 14 to 16 (Fig. 3). The number of nodes in the hidden layer was set to 3, resulting in a compact 21-3-1 network architecture. Alternative architectures with larger number of hidden nodes were evaluated [Atzberger and Rembold, 2009] but resulted in overfitting problems.

**Table 2 - Overview over the experiment investigating the portability of trained networks to forthcoming growing seasons in an operational setting.** Training is based on a stratified sub-sampling within data set 1, whereas the evaluation was done on data set 2.

<table>
<thead>
<tr>
<th>Aim</th>
<th>Training data</th>
<th>Sample generation</th>
<th>Sub-setting</th>
<th>Number of models</th>
<th>Evaluated variable</th>
<th>Validation data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portability across years</td>
<td>Data set 1</td>
<td>Stratified sampling</td>
<td>3/8</td>
<td>16</td>
<td>Total winter crop acreage of Tuscany (∑AFIw)</td>
<td>Data set 2</td>
</tr>
<tr>
<td>(n=8 yr)</td>
<td>(n=8 yr)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(n=8 yr)</td>
</tr>
</tbody>
</table>


For network training the resilient backpropagation algorithm was used within Matlab programming environment [The Mathworks Inc., 1994-2011]. To achieve good generalisation properties of the net-work, the early stopping technique [Demuth and Beale, 2003] was applied. The training samples were split into three subsets with 50 % (training), 25 % (test) and 25 % (validation) of the total available pattern. The training set was used for computing the gradient and up-dating the weights and biases. The error on the test set
was monitored during the training process. When the network began overfitting the data, the test error started to rise. At this point the training was stopped automatically and the actual weights at the minimum of the test error were returned. Nets trained in this way were subsequently used for estimating the target variable of the training and validation datasets. The experiment described in this paper is summarized in Table 2: sixteen (3-year) subsets of the 8-year reference data set were used for network training, yielding 16 different models. The trained nets were next applied to the time series (2002-2009) to derive regional estimates of winter crop acreage. The regional estimates were compared to the official (AGRIT) statistics of the region.

The sixteen (3-year) subsets were chosen so that always data from contrasting winter crop conditions were included (from low to high). For this purpose, the eight years were split into three categories: low winter crop surface (2 years – L: < 140.000 ha), medium (4 years – M: 140.000-180.000 ha), and high winter crop surfaces (2 years – H: >180.000 ha) (Fig. 4). This stratification results in sixteen (three of eight year) combinations: 2 L x 4 M x 2 H (Fig. 8).
Results
Sixteen nets were trained with reference data 1988-2001 (data set 1) for predicting the total winter crop acreage of the region of Tuscany for the time period 2002-2009 (data set 2). For each net, a different combination of 3 years of available reference data (≤ 2001) was used for training. To select the training samples, a stratified sampling was chosen always ensuring inclusion of training data from growing seasons with low (L) to high (H) winter crop acreages (Fig. 4 and Fig. 8).

The time profile of predicted winter crop acreages resulting from the sixteen nets are shown in Figure 9 (top, in green) together with the official statistics (in red). A scatterplot of observed (AGRIT) and modelled surfaces (mean and standard deviation) is shown in Figure 9 (bottom). The 1-to-1 line is added for orientation. The curves in Figure 9 (top) testify a considerable variability in the cropped surfaces (red line), mainly as a direct result of European agricultural policies and market prices. The main features of the official statistics are relatively well reproduced by the network approach (in green). On average (not shown), about half (47 %) of the total variance in the reference data was accounted by individual nets (%RMSE around 18 percent corresponding to 28,000 ha in absolute values; NSE: 0.46). The predictions of individual nets varied stronger for medium high surfaces (Fig. 9; bottom) and increased towards later years (Fig. 9; top). The prediction power of individual nets varied considerably in terms of $R^2$, RMSE, %RMSE and NSE (Fig. 10), mainly because of the simulations for 2008, 2007 and 2005. As only data from data set 1 was used for network calibration (1988-2001) it looks plausible that network...
outputs were more variable for (later) years. Also one has to keep in mind that despite the fact
that the training data were chosen through stratified sampling (Fig. 8), the nets received training
data with different representativeness and characteristics (Fig. 4). The observed variability of
$R^2$, RMSE, %RMSE and NSE values for individual nets shown in Figure 10 is a direct result of
this sampling effect.

Figure 9 - Estimated total winter crop acreage for Tuscany obtained
with 16 neural nets trained with 3 out of 8 years of available reference
data (data set 1) and applied to the time series (2002-2009) (data set
2). From each of the sixteen nets, the estimated winter crop acreage
is obtained for the period 2002-2009. (top) The green area indicates
the variability resulting from the individual nets. The reference data
(official AGRIT statistics) are shown in red. (bottom) Official crop
areas vs. estimates. The boxes and whiskers indicate the simulated
averages and standard deviations across the sixteen individual nets.
The 1-to-1 line is also drawn.
This interpretation is supported by the fact that the best performing net (results presented in Fig. 11) was trained with data including the two most extreme years of data set 1: the year with the lowest overall winter crop area (1992) and the year with the highest area (1991). Using ‘extreme’ data enables the net to better interpolate and reduces the need for extrapolation. However, in practice it may be difficult to obtain the necessary data covering the full range of crop acreages.
For the particular net shown in Figure 11, a coefficient of determination of 0.68 was achieved with a RMSE around 23.500 ha (roughly 15 percent of the average) with NSE of 0.61. The estimates group relatively well around the 1-to-1 line, with only little bias (observed average: 155.710 ha; simulated average: 152.680 ha). Important scatter towards the two extremes and a slope < 1 can be observed mainly resulting from the fact that the training data (data set 1) did not cover the full range of variation in the evaluation data set (data set 2). In such a case, the network has to extrapolate, as it is confronted with signatures not seen previously during the training phase. The results underpin the importance of high-quality training information for model calibration [Chang et al., 2007].

Discussion

The objective of this research was to develop and validate a network-based unmixing approach (1) for mapping the annual winter crop distribution at sub-pixel scale and (2) for deriving total crop acreages for the region of Tuscany using coarse resolution AVHRR NDVI 1 km time series. Encouraging results were obtained despite an environmentally complex and diverse study area and a limited amount of input data (e.g., Landsat imagery, AVHRR time series, CORINE LC map and regional agricultural statistics). The success relates partly to the compact network architecture selected for the study. Indeed, only three nodes were used in the hidden layer, albeit training samples could be predicted with higher accuracy if more complex networks were used [Atzberger and Rembold, 2009]. The compact network guarantees robustness of the mapping technique as overfitting and overspecialization problems increase with the number of nodes in the hidden layer. With only three nodes in the hidden layer and receiving additional information regarding the fractional coverage of the primary land cover classes (e.g., forest, plantation, etc.) a good compromise was found between network flexibility and robustness.

The robustness and portability of the trained networks was testified when applied to the AVHRR time series (2002-2009) while training data was restricted to the time period before 2002. The main changes in cropped areas, mainly driven by political decisions, market prices and weather-related management decisions, were relatively well reproduced by the net. For example the introduction of uncoupled payments in 2005 lead to a clear wheat area decrease culminating in the historical minimum of 2006 and after having reached an absolute maximum in 2004, the last year of direct payments. From 2006 when the reform effects can be considered stable, the wheat area has grown again thanks to favorable wheat prices, while the 2009 decline is due to a combined effect of unfavorable weather conditions (rainy sowing season in 2008) and less attractive prices of wheat as compared to other crops like soya, sunflower and rapeseed [AGRIT, 2009]. While not all features were equally well recognized by the proposed approach a general agreement can be noted.

Past studies demonstrated already the potential of using NDVI time series to study vegetation dynamics and (land cover) changes. The importance of image temporal frequency for accurately detecting forest changes was documented by Lunetta et al. [2004]. The utility of NDVI time series is limited, however, by the availability of high-quality (e.g., cloud free) data [Lunetta et al., 2006]. To cope with this critical factor, the AVHRR NDVI 1km data used in the current study were pre-processed using a method developed by Eilers [2003] and Atzberger and Eilers [2011]. Pre-processing was conducted to provide a filtered (anomalous data removed), cleaned (excluded and missing values estimated) and smoothed
uninterrupted data stream supporting the neural networking. The quality of the selected smoother for filtering remotely sensed time series was checked in Atzberger and Eilers [2011] and recently in Atkinson et al. [2012]. The studies found that the selected smoother yields internally consistent data well representing the phenology of various vegetation cover types. In the conference paper of Atzberger and Rembold [2009] using the same data as for the current study, a positive effect of the filtering was noted when compared to results using the original (unfiltered) AVHRR data.

A shortcoming of the proposed approach relates to the broad class of “summer” crops within the arable land, which may show a high inter-annual (and spatial) variability. This could be clearly seen both in the inner and coastal plains of Tuscany where it is known that the proportion, composition and phenology of summer crops have a high inter-annual variability. In the present study, negative effects related to the mentioned variability of temporal “summer” crop signatures were minimized by excluding time stamps in the NDVI time series where winter and summer crops strongly overlap. If feasible, one should try to derive and use more detailed information regarding the spatial distribution of different summer crops within the study area.

The study did not assess the utility of time series derived from radiometrically improved sensors such as MODIS or SPOT VGT. Possibly, results could be further improved. Such improvements would come at the cost, however, of much shorter time series. Only by using data from NOAA AVHRR, it is possible to globally monitor cropping pattern back to the eighties of the last century. The study demonstrated that it is possible to base such an analysis on internally consistent NDVI profiles from AVHRR starting in 1988 if a suitable filter is applied to smooth the data prior to data analysis.

Several studies pointed out that the primary limitation for crop-type mapping is not the input data or the employed un-mixing algorithm, but the presence of high-quality training information for model calibration [e.g. Chang et al., 2007]. In this respect is has to be mentioned that the current study did not evaluate the quality of the high resolution reference information derived from official AGRIT statistics and Landsat TM/ETM+ data. Implicitly it was assumed that the spatialization of the regional (AGRIT) statistic – and the statistic itself – were accurate. Possible errors in the reference data would negatively affect network training and validation. To overcome this problem ground truth information should be acquired. This was however beyond the scope of the present study.

Conclusions

The study evaluated the performance of neural networks for estimating the inter-annual variation in regional winter crop acreage using low resolution imagery and ancillary information. Reference information concerning the sub-pixel crop acreages was derived from official regional statistics using a simple thresholding approach applied to high resolution Landsat TM/ETM+ imagery. For assessing the robustness and portability of the developed tool, an independent statistical validation approach was chosen. In this validation, the estimated total winter crop area for the entire study region (Tuscany) was compared against official (AGRIT) statistics, using data not previously used for network training.

Several important conclusions can be drawn from this study:

- From our experience, we recommend compact networks with only a few nodes in the hidden layer (e.g. 3 nodes);
The net correctly predicted at least half of the inter-annual variance in the regional winter crop acreage (normalized RMSE of around 15-20 percent) when reference information from three contrasting growing season was made available for network training. The results were better than those obtained previously with the same data set using spectral angular mapping (SAM) of temporal NDVI profiles [Rembold and Maselli, 2006];

Our findings reveal the importance of using a non-linear mapping technique (such as neural nets) able to include possibly available ancillary information (here CORINE land cover information);

With the best performing net (trained with data including the two most ‘extreme’ years of the training data base) up to 68 percent of the inter-annual crop area variability of Tuscany was explained. This highlights the need for high-quality training data. The term ‘high-quality’ implies in this context (i) the accuracy of the reference information and input data, but also (ii) how well the reference data covers the full range of crop area dynamics in the study region;

Albeit the performance of the neural nets decreased when applied to different growing seasons, the drop in accuracy was not as strong as found for example by Bossyns et al. [2007].

The strength of the neural net is its speed in processing large amounts of data and its capability to integrate remotely sensed time series and ancillary information derived from existing land cover maps. The use of ancillary information allows the net to better distinguish NDVI time series from pixels with similar proportions of winter crops and arable land but with different “background” contributions (e.g. various proportions of forests, pastures and plantations within the non-arable land). Other input data sets can easily be adapted to operational settings.

As for any other modelling tool, the quality of the provided reference information is of utmost importance in neural networking. In the optimum case, data from growing seasons covering the most “extreme” cropping pattern should be made available for network training. This avoids later confronting the network with input vectors not seen during the training step.

Considering the results obtained, it would be of interest translating the approach to other geographical areas (both important crop production areas and food insecure countries) and to evaluate other low or medium resolution NDVI time series such as SPOT VGT or MODIS. The main requirements for further investigations in this sense are the availability of (1) training data for the discrimination between winter and summer crops, either based on high resolution data analysis or using existing agricultural databases (e.g. high resolution land use classifications or cadastral data) [Verbeiren et al. 2008], and (2) reliable crop area statistics, possibly based on a high accuracy area frame sampling approach, such as the one of the AGRIT project at provincial or national level. For situations where these criteria are met, the methodology proposed in this paper could be exploited for an operational crop area change estimation system. As shown in Figure 3, the method uses coarse resolution imagery up to dekad 25 (beginning of September) which means that winter crop acreages in Tuscany and in areas with similar agro-climatic conditions can be estimated already by end of September.

This may be an advantage for those users that need crop-specific acreage information – or at least some broad trends – well before the release of official statistics, which usually (in Europe) are not available before the following year.
Acknowledgements
The authors thank the anonymous reviewers for their useful and constructive comments. Thanks to those having made the AGRIT statistics available for this study.

References


Received 13/10/2011, accepted 26/01/2012