

Article

# Nonlinear Spectral Unmixing for the Characterisation of Volcanic Surface Deposit and Airborne Plumes from Remote Sensing Imagery

Giorgio A. Licciardi <sup>1,2,\*</sup>, Pasquale Sellitto <sup>3</sup>, Alessandro Piscini <sup>4</sup> and Jocelyn Chanussot <sup>2,5</sup>

<sup>1</sup> Research Consortium Hypatia, 00133 Rome, Italy

<sup>2</sup> GIPSA-Lab, Institut Polytechnique de Grenoble, 38000 Grenoble, France;  
jocelyn.chanussot@gipsa-lab.grenoble-inp.fr

<sup>3</sup> Laboratoire de Météorologie Dynamique, Institut Pierre Simon Laplace, École Normale Supérieure, 75231 Paris, France; psellitto@lmd.ens.fr

<sup>4</sup> Istituto Nazionale di Geofisica e Vulcanologia, 00143 Rome, Italy; alessandro.piscini@ingv.it

<sup>5</sup> Faculty of Electrical and Computer Engineering, University of Iceland, 101 Reykjavik, Iceland

\* Correspondence: giorgio.licciardi@consorzioipazia.it; Tel.: +39-06-8567696

Received: 29 March 2017; Accepted: 20 June 2017; Published: 23 June 2017

**Abstract:** In image processing, it is commonly assumed that the model ruling spectral mixture in a given hyperspectral pixel is linear. However, in many real life cases, the different objects and materials determining the observed spectral signatures overlap in the same scene, resulting in nonlinear mixture. This is particularly evident in volcanoes-related imagery, where both airborne plumes of effluents and surface deposit of volcanic ejecta can be mixed in the same observation line of sight. To tackle this intrinsic complexity, in this paper, we perform a pilot test using Nonlinear Principal Component Analysis (NLPCA) as a nonlinear transformation, that projects a hyperspectral image onto a reduced-dimensionality feature space. The use of NLPCA is twofold: (1) it is used to reduce the dimensionality of the original spectral data and (2) it performs a linearization of the information, thus allowing the effective use of successive linear approaches for spectral unmixing. The proposed method has been tested on two different hyperspectral datasets, dealing with active volcanoes at the time of the observation. The dimensionality of the spectroscopic problem is reduced of up to 95% (ratio of the elements of compressed nonlinear vectors and initial spectral inputs), by the use of NLPCA. The selective use of an atmospheric correction pre-processing is applied, demonstrating how individual plume and volcanic surface deposit components can be discriminated, paving the way to future application of this method.

**Keywords:** nonlinear spectral unmixing; nonlinear PCA; volcanic plumes; hyperspectral remote sensing

## 1. Introduction

Volcanoes can inject a great amount of gaseous and particulate effluents to the atmosphere, like water vapour, sulphur dioxide and ash, both at background degassing conditions and during explosive eruptions [1]. Depending on the volcanic activity and the emission fluxes, these effluents can organize as volcanic clouds, which interact with solar and terrestrial radiation, thus affecting the observed spectra in the remotely observed images. The main types of volcanic aerosols are mineral ash, directly emitted by the eruption and secondary sulphate aerosols (SSA), which are produced in-plume by oxidation and hydration of SO<sub>2</sub> emission. Ash can be detected by thermal infrared (TIR) observations using absorption signatures between 8 and 12 μm (1250 and 833 cm<sup>-1</sup>), typically centred around 10 μm (1000 cm<sup>-1</sup>) (e.g., [2,3]). Sulphur dioxide emissions are measured both with TIR (e.g., [4–6]) and ultraviolet/visible (UV/VIS) satellite instruments [7]. Recently, targeted sensitivity analyses have

shown that also SSA have a typical spectral signature in the TIR spectral range that can be used to detect and characterise this volcanic particulate product [8,9].

Apart from quantitative and semi-quantitative observation of these emitted species using high spectral resolution (sounder) instruments, the interaction of the radiation with the volcanic plumes allow the detection and tracking of the volcanic plume and the identification of the species therein, by multi- and hyper-spectral (imaging) observations (e.g., [10–14]). Recently, neural network (NN) approaches have been tried in order to quantify volcanic ash and SO<sub>2</sub> using both multi-spectral and hyper-spectral data [15–17]. In such cases, unmixing the volcanic cloud spectral signature from other spectral features, as those arising from other atmospheric components in the line of sight of the instruments or from the surface, is vital to rule out these different contributions [3,18].

Hyper-spectral imaging sensors normally record scenes in which numerous interacting objects and material substances, both at the surface and in the overlying atmosphere, contribute to the spectrum measured from a single pixel, by their interaction with the atmospheric radiation recorded by the sensor. Given such mixed pixels, the process of identification of the individual constituent materials in the mixture (endmembers), as well as the proportions in which they appear (abundances), is commonly referred to as spectral unmixing. In remote sensing images, usually the endmembers correspond to the spectral response of macroscopic materials present in the scene, such as surface water, soil, human structures (like buildings) and dominating atmospheric features (like thick meteorological or aerosols clouds) [19].

In recent literature, unmixing techniques are characterized as linear or nonlinear processes [20]. Linear mixtures are dominant when the incident light interacts with objects composed of one individual material before reaching the sensor (different physical/micro-physical properties but the same chemical/mineralogical composition). Thus, light reflected from different materials are mixed in the sensor itself, with minimal interference [21–23]. On the other hand, nonlinear mixing occurs as the result of physical interactions between light scattered from multiple materials, with different chemical/mineralogical composition. These interactions can be sub-divided into *multi-layered* and *microscopic* interactions. Multi-layered interactions occur when light, reflected from one individual material, interacts with other individual and distinct objects before reaching the sensor. Microscopic mixing occurs when two or more materials are physically mixed and this mixture interacts with radiation. In the case of the multi-layered mixing, the first order terms are sufficient to describe the mixture leading to the bilinear model [24,25]. On the other hand, microscopic interactions require an extremely complex physical modelling of the mixture and the interaction with radiation. For these reasons, only approximation are presently proposed in literature [26,27]. However, techniques developed for the processing of internal mixtures are inefficient in the multiple interaction scenario (and vice versa). Moreover, these models are based on the assumption that the mixtures are produced by different materials having Lambertian surfaces. In many real cases, the light that interacts with different materials does not produce in an isotropically distributed radiance. All these effects result in nonlinear mixing.

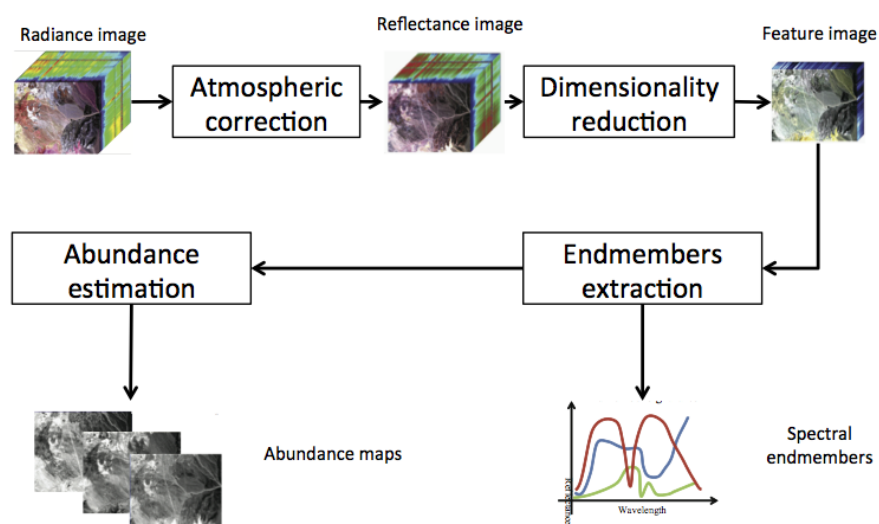
A more flexible solution can be achieved by using machine learning approaches, such as NN [28–30]. These algorithms can learn nonlinear correlations in a supervised fashion based on a collection of examples (training dataset). However, the universal approximation properties of NN [31] assumes that the training dataset covers all the possible physical scenarios and interactions between radiation and fixed materials. Practically, a very large training dataset is necessary to approach the convergence of NN properties in several possible scenarios, given the problem under investigation. This large training dataset is not always available. A possible alternative is to project the hyper-spectral image into a linearised feature space by means of a nonlinear transformation. In this way, any kind of linear unmixing method can be effectively applied to these linear features. Based on this idea, in this paper, we propose the use of the nonlinear Principal Component Analysis (NLPCA) for the projection of the hyper-spectral image into a feature space [32]. The obtained features are then used as input to a linear

unmixing algorithm for the identification of the endmembers. Finally, the abundance estimation is performed solving the constrained least square problem.

The remainder of the paper is organized as follows: in Section 2, the the proposed approach is introduced, while in Section 3, experimental methodology as well as an in-deep analysis of the results are provided. Finally, Section 4 gives some conclusion remarks.

## 2. Nonlinear Spectral Unmixing

The general framework of spectral unmixing, either linear or nonlinear, is composed of three main processing steps: (1) dimensionality reduction; (2) endmember identification and (3) abundance estimation, as depicted in Figure 1. In cases where the surface characterisation is the sole target of the analysis, the observed radiances can be converted to surface reflectances by means of an atmospheric correction, aimed at compensating for the atmospheric attenuation and scattering. Performing an accurate atmospheric correction is often an arduous task, due to coarse information on the atmospheric composition. Moreover, applying a correction for the atmospheric contribution may have a negative impact on the spectral unmixing when the task of the analysis is to find endmembers present in the atmosphere. The three processing steps are discussed in the following.



**Figure 1.** Schematic diagram describing the complete processing for hyperspectral unmixing.

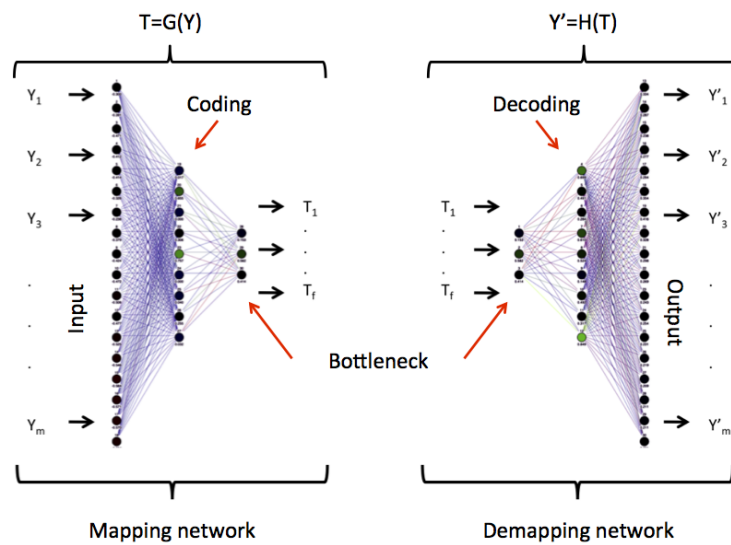
As for step 1 (dimensionality reduction), since hyperspectral images are composed by hundreds of extremely correlated bands, it is possible, and indeed beneficial, to reduce the effective dimension of the input data by use of decorrelation approaches. This processing step projects the image into a reduced dimensionality feature space. The algorithm performance in terms of computation time, complexity and performances are, then, generally improved [33]. However, a careful selection of the dimensionality of the feature space is essential because this choice limits the number of possible final endmembers. As for step 2 (endmembers identification), there exist several approaches to rule out endmembers in a given image scene, which fall into three main groups. Geometrical approaches are based on the hypothesis that linearly mixed vectors are in a simplex set. Statistical approaches use parameter estimation techniques to determine the endmember. Sparse regression approaches formulate unmixing as a linear sparse regression problem. As for step 3 (abundances estimation), given the hyperspectral image and the endmembers identified in the previous step, the abundance of each endmember is quantified solving a constrained optimization problem, which minimizes the residual between the observed spectral vectors and the linear space spanned by the inferred spectral signatures. Usually, in linear unmixing, the fractional abundances are constrained to be nonnegative and to sum to one. While atmospheric correction and dimensionality reduction are optional, the

endmembers' determination and abundances' estimation steps are central to any unmixing approaches. Often, steps 2 and 3 are implemented simultaneously [34]. In the case of nonlinear mixtures, the use of linear approaches may result in the false identification of non-physical endmembers and in most cases cannot detect the whole of endmembers actually present in the scene. In the specific case of nonlinear approaches, the endmembers' determination and abundances' estimation steps require a prior detailed knowledge of the possible physical interactions between materials, which is not always available. In order to override this limitation, in this paper, we propose a novel approach that, instead of deriving an extremely complex nonlinear model, linearizes the input data by means of a preliminary nonlinear transformation. In particular, we propose to use the NLPCA to project the original hyperspectral image into a linearized feature space. The advantage of the use of NLPCA is twofold. On the one hand, it performs a linearization of the original information. On the other hand, it reduces the data dimensionality. The NLPCA, thanks to its nonlinear functions basis, permits obtaining a set of features that do not present nonlinear correlations, thus allowing the subsequent use of linear approaches for the endmember determination and inversion phases. As for the endmember identification, we proposed the use of the N-FINDR algorithm [35]. This choice is based on the main advantage of the N-FINDR, which is able to automatically detect endmembers without a priori information. This characteristic is extremely valuable in the case of NLPCA, where the obtained features may not always have a physical interpretation. The abundance estimation step will be then carried out by using the SUnSAL algorithm [36] for the solution of the constrained least square problem. In the following, a detailed description of the NLPCA approach, as well as the N-FINDR and SUnSAL algorithms are reported.

### 2.1. Nonlinear Principal Component Analysis

In this work, NNPCA, commonly referred to as a nonlinear generalization of the PCA techniques, is performed by an AutoAssociative Neural Network (AANN) or auto-encoder [32]. A standard auto-encoder is a conventional feedforward neural network, having a symmetrical three layer topology, where input and output layers have the same number of nodes, and a hidden layer, usually referred to as bottleneck, of smaller dimension than either input/output layers. Both input and output layers have sigmoidal activation functions. The auto-encoder is trained to perform identity mapping, meaning that the input has to be equal to the output [37]. This means that, after a successful training phase, the fewer nodes in the bottleneck layer, than in the input/output, represent (in fact encode) the information of the inputs in a smaller dimensionality space. Due to the nonlinear nature of NNs, this information compression is obtained with nonlinear combinations of the inputs. In other words, data compression caused by the network bottleneck forces hidden units to represent significant features in the data, removing redundancies. The smaller-dimensionality information vector can, then, be used as input for the subsequent processing.

Different from standard auto-encoders, the topology of a nonlinear AANN uses by default three hidden layers, including the internal bottleneck layer of smaller dimension than either input or output layers (Figure 2). In order to understand why three hidden layers are necessary to obtain a nonlinear representation of the data, it is useful to consider the nonlinear AANN as a combination of two successive neural networks or functional mappings, namely *coding* and *decoding sub-networks*, as depicted in Figure 2.



**Figure 2.** General structure of the Autoassociative Neural Network (AANN). The AANN can be seen as a concatenation of two simple networks. The first network project the original data  $Y$  into a lower dimensionality subspace  $T$  by means of the transformation function  $G$ . The second network reproject the compressed data  $T$  back into the original space by means of the transformation function  $H$ .

The first sub-network represents the encoding or extraction function:

$$T = G(Y), \tag{1}$$

which projects the original  $m$ -dimensional data  $X$  onto a lower dimensional subspace defined by the activations of the units in the central hidden layer (bottleneck), producing a smaller-dimensionality (say  $n$ -dimensional) vector  $T$  by means of the mapping  $F$ . Similarly, the second sub-network defines an arbitrary functional mapping:

$$Y' = H(T), \tag{2}$$

which projects from the smaller-dimensionality feature space back onto the original  $m$ -dimensional space, by means of the mapping function  $H$ . The ability of any NN to fit arbitrary nonlinear functions depends on the presence of a hidden layer with nonlinear nodes. In [31], it was shown that any nonlinear function can be approximated by a superposition of a set of  $\sigma(x)$  transformations that are continuous, bounded and monotonically increasing functions, with  $\sigma(x) \rightarrow 1$  as  $x \rightarrow +\infty$  and  $\sigma(x) \rightarrow 0$  as  $x \rightarrow -\infty$ . This property is often called *universal fitting* [38], and is a generalisation of the Weierstrass theorem (which applies to polynomial functions). Thus, a network lacking a hidden layer is only capable of producing linear combinations of the inputs, given linear nodes in the output layer. In the same way, a network lacking a hidden layer but including nonlinear activation functions in the output layer is only capable of approximating multi-variable sigmoidal functions. Similarly, a NN with linear nodes in the hidden layer will return linear combinations of the inputs. From these considerations, it can be affirmed that the NLPCA can be implemented by two NNs approximating the nonlinear functions  $G$  and  $H$ . The NN producing the  $G$  mapping has as an input layer of  $m$  nodes followed by the hidden layer (often called *mapping layer*) with  $m_1 > n$  nodes and sigmoidal transfer functions (to assure the *universal fitting* property of NNs). The output layer of this subnet contains  $n < m$  nodes and for this reason it is often called *bottleneck*. The second NN (also called *decoding subnet*) producing the  $H$  mapping, has an input layer with  $n$  nodes, followed by the hidden layer (often called *demapping layer*) with  $m_2 > n$  nodes and sigmoidal transfer functions. The output layer yields the reconstructed data and thus contains  $m$  nodes.

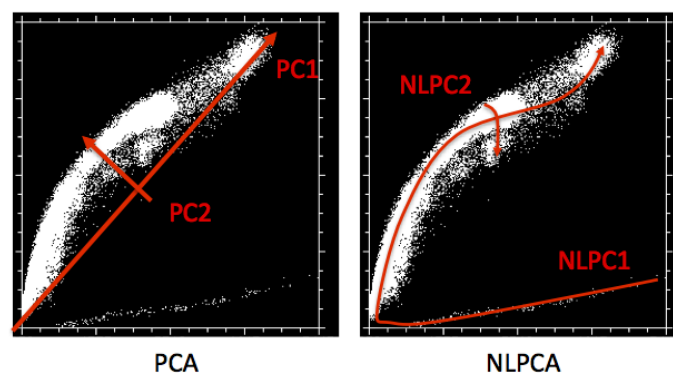
In order to correctly define the two mapping functions  $G$  and  $H$ , a supervised training of the two NNs is required. This needs, in principle, a complete knowledge of the relations between the input and output spaces, which is only an ideal case. From a practical point of view, this means that while the inputs of the coding subnet is known, the output is unknown. Conversely, the input of the decoding layer is unknown while its output is known (i.e., the input, as the coding+decoding NN are aimed at approximating an identity mapping). Therefore, a direct supervised training of the two sub-networks is not feasible. However, one can observe that combining in series the two ANNs is equivalent to define a composite function  $Y' = H(G(Y))$  that links the original data  $Y$  with its reconstruction version  $Y'$ . Thus, practically, the combined network is trained to produce the identity mapping. This means that the parameters of the network representing  $Y'$  are optimized so that the reconstructed outputs match the inputs as closely as possible. The training aimed at learning the identity mapping has been called *self-supervised backpropagation* or *autoassociation*, leading to the definition of these NNs as AutoAssociative NNs (AANNs).

For AANNs, the training phase is an iterative process and is completed when the following sum of squared errors is minimized:

$$E = \sum_{p=1}^n \sum_{i=1}^m (\hat{y}_i - y_i)_p^2. \quad (3)$$

In Equation (3),  $\hat{y}_i$  and  $y_i$  are the calculated and target output vectors of the AANN, for a training set of  $p$  examples.

Once the AANN is trained, it is possible to use extract the *coding sub-network* only, to project the original data into a lower dimensional space given by the bottleneck layer. Thus, the  $f$  NLPCA can be obtained from the bottleneck. From a topological point of view, it can be noted that, while the number of nodes in the bottleneck layer defines the features subspace, the nodes in the coding and decoding layers are related to the complexity of the mapping and demapping functions. As it can be seen in Figure 3, the NLPCA defines a set of nonlinear functions able to describe the nonlinear correlations between input variables. The result is a linearised feature space. However, one of the main difficulties in designing the AANN relies in the selection of the correct number of nodes in the three hidden layers, since the mapping functions as well as the subspace dimension strongly depend on them.



**Figure 3.** Example of a nonlinear dataset being mapped into feature space by means of linear and nonlinear PCA.

The best NN topology can be retrieved by using a simple heuristic grid search algorithm that varies recursively the number of nodes of the hidden layers and evaluated the value of the Means Square Error (MSE) error [39]. Then, the topology presenting the smallest error is selected. However, without starting assumptions, this approach can be extremely time consuming and different optimal solutions should be found. Starting from  $p$  that represents the number of samples in the training set, a separate constraint is imposed by each output node, so that the total number of the possible adjustable parameters (weights and biases for all network connections and nodes, respectively) must be less than

$p \times m$ . Moreover,  $m$ ,  $M_1$ ,  $M_2$  and  $n$  being the number of nodes of input/output, coding, decoding and bottleneck layers, and analysing the structure of the AANN used here, it can be found that the number of adjustable parameters is  $(M_1 + M_2)(m + n + 1) + m + n$  that implies the following inequality:

$$M_1 + M_2 \ll \frac{m(n - 1) - n}{m + n + 1}. \tag{4}$$

The aim of a dimensionality reduction method is to reduce the original spectral dimension into a lower dimensional space. This can be translated into the AANN structure as a condition on  $n$ , i.e.,  $n \ll m, p$ . Then, Equation (4) becomes:

$$M_1 + M_2 \ll n. \tag{5}$$

Assuming a balanced structure of the AANN,  $M_1$  and  $M_2$  should have the same dimensions ( $M_1 \sim M_2 = M$ ), we have:

$$2M \ll n. \tag{6}$$

It is worth noting that Equation (6) is effective only if the number of mapping/demapping nodes  $M$  is greater than the number of nodes in the bottleneck layer  $n$ . Otherwise, there will not be enough data to effectively extract  $n$  NLPCs. It is also worth underlining that, since the output has to simply replicate the input, there is no need to have a specific a priori knowledge for the learning phase implementation. This implies that the AANN training can be performed in a fully automatic way and that all pixels in the image can be used as training samples for this task. Practically, this has actually been the technique adopted in this paper.

### 2.2. Endmember Extraction and Abundance Estimation

After the dimensionality reduction phase been accomplished, the subsequent endmember extraction and abundance estimation phases need to be accomplished. Among the several algorithms developed for automatic or semiautomatic extraction of spectral endmembers, the N-FINDR algorithm attempts to automatically find the simplex of maximum volume that can be inscribed within the hyperspectral data set [35]. The N-FINDR implementation is firstly initialized by randomly selecting a set of  $q$  endmembers  $\{E_1, E_2, \dots, E_q\}$ , where  $q \leq n + 1$ , and  $n$  corresponds to the dimension of the feature space. Then, the volume of the simplex defined by the current set of endmembers is derived by:

$$V(E_1, E_2, \dots, E_q) = \frac{\left| \det \begin{bmatrix} 1 & 1 & \dots & 1 \\ (E_1 & E_2 & \dots & E_q) \end{bmatrix} \right|}{(q - 1)!}. \tag{7}$$

Then, for each pixel vector  $X(i, j)$  of the input hyperspectral data, the volume  $V$  is recalculated by testing the pixel in the first endmembers position:

$$\begin{aligned} &V(X(1, 1), E_2, \dots, E_q) \\ &V(X(1, 2), E_2, \dots, E_q) \\ &\vdots \\ &V(X(r, c), E_2, \dots, E_q), \end{aligned} \tag{8}$$

where  $r$  and  $c$  represent the number of rows and columns of the image. If one of the volumes calculated in Equation (8) is greater than  $V(E_1, E_2, \dots, E_q)$ , then  $E_1$  is replaced with the pixel corresponding the the maximum volume, and a new set of endmembers is produced. The same procedure is then carried out iteratively by testing the volumes in the other endmembers positions, retaining the combinations corresponding to the maximum volumes. The processing ends when all the pixels in the input data have been tested in each endmember position.

Since in the proposed approach the input data corresponds to the features derived by the NLPCA, the obtained endmembers are represented in the feature space. In order to obtain the equivalent endmembers in the spectral domain, the decoding sub-network of the NLPCA is used.

Once the endmembers are obtained, the fractional abundances estimated by minimizing the total squared error, under the constraints of non-negativity and/or the sum to one:

$$\min_x \left( \frac{1}{2} \right) \| Ea - S \|^2, \tag{9}$$

$$x \geq 0,$$

$$1^T x = 1,$$

where  $E \in \mathbb{R}^q$  denote the matrix containing the  $q$  endmembers,  $a \in \mathbb{R}^k$  the fractional abundance vector, and  $S \in \mathbb{R}^k$  the observed mixed pixel. In order to solve the optimization problems, we used the SUnSAL algorithm, which is an instance of the Constrained Split Augmented Lagrangian Shrinkage Algorithm (C-SALSA) methodology to effectively solve a large number of constrained least-squares problems sharing the same matrix system in [36]. Figure 4 depicts the complete schematic of the proposed approach.

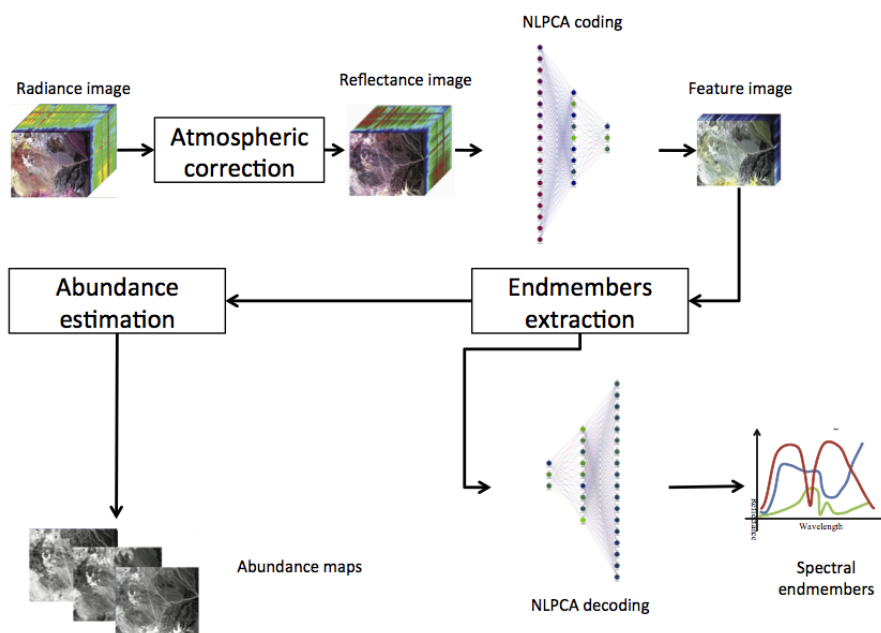


Figure 4. Schematic diagram describing the proposed processing for hyperspectral unmixing.

### 3. Experimental Results

In this section, two real measurements datasets have been considered to test the proposed technique. Both radiance and reflectance data, i.e., without and with atmospheric correction, are used in order to analyze the effect of these two options in our approach. Differently from simulated data, real images are, generally, strongly conditioned by a great deal of additional circumstances, such as differences in illumination through the scene, angle of view as well as multiple scattering effects. These factors have been reported to influence the endmember selection in linear unmixing approaches [34]. As for the assessment of the effectiveness of the proposed technique, a qualitative analysis has been carried out by considering comparisons with ground truth fractional abundance maps and spectral library. In the case of abundance maps’ ground truth, the effectiveness of the technique can be assessed by estimating the abundances of the endmembers in the scene and comparing the obtained values with reference fractions. In using ground truth spectral library, the quality of the endmembers is evaluated by comparing them with some reference spectral signatures using spectral similarity



criteria. In order to further appreciate the effectiveness of the proposed method in handling nonlinear mixtures, the obtained results have been compared with those obtained applying the endmembers extraction and abundance estimation phases, directly to the hyperspectral image and to the linear features obtained through the use of standard PCA.

### 3.1. Campi Flegrei

On a first experiment, we applied the proposed technique to a hyperspectral image acquired with the Hyperion satellite-borne sensor, onboard the EO-1 satellite. Hyperion acquires 220 hyperspectral bands in the Visible/near-infrared, from 0.4 to 2.5  $\mu\text{m}$ . However, only 155 bands have been retained from the original dataset, discarding the most noisy and the bands without relevant information for this application [40]. The considered Hyperion image has been acquired in 2008 over the Campi Flegrei (CF) area, northwest of Naples, Italy. In this experiment, we focused on the caldera area, comprising more than 24 craters and volcanic edifices and presenting effusive gaseous manifestations, in particular in the Solfatara crater.

The CF region is located in the Campanian plain, which is a NW-SE trending Plio-Quaternary extensional basin bordered by carbonate platforms. The CF caldera has been interested by volcanism and hydrothermal activity for thousands of years [41]. The main structures of CF consist of two nested calderas formed after the Campanian Ignimbrite and the Neapolitan Yellow Tuff eruptions [42]. The outer is the Campania Ignimbrite (CI) caldera, dated 39 ky, while the inner is the Neapolitan Yellow Tuff (NYT) caldera, dated 15 ky [43]. Currently, the surroundings of Solfatara crater are the area with the strongest geothermal emission in CF. Large quantities of volcanic-hydrothermal  $\text{CO}_2$  are released through soil diffuse emission. The gas emissions from this area consists in about 5000  $\text{t day}^{-1}$  of a  $\text{CO}_2/\text{H}_2\text{O}$  mixture. Its power is 100 MW, which is 10 times higher than the conductive heat flux over the whole caldera surface [44]. Since the image has been acquired from satellite, the atmospheric contribution has a relevant role. In this first example, an atmospheric correction is applied. Figure 5 reports the ground truth of the Hyperion image.

The Hyperion image after the atmospheric correction has been processed to mitigate the striping effect. The corrected image has been then projected in the feature space by means of the NLPCA. In particular, we found that the optimal topology for the AANN was 155 input/output nodes, 80 nodes in the coding/decoding layers and nine nodes in the bottleneck layers. Then, the dimensionality of the hyperspectral input vector is the reduced by about the 95% by the nonlinear compression of the implemented AANN.

Once the AANN is trained, the N-FINDR algorithm has been applied to the nine nonlinear principal components obtained from the bottleneck layer of the AANN and 10 endmembers have been detected. The obtained endmembers have been used as input to the SUnSAL algorithm in order to estimate the fractional abundances of each endmember (Figure 6). Finally, the 10 endmembers (EM, in the following) have been processed through the decoding sub-network of the AANN in order to retrieve their spectral signatures reported in Figure 7.

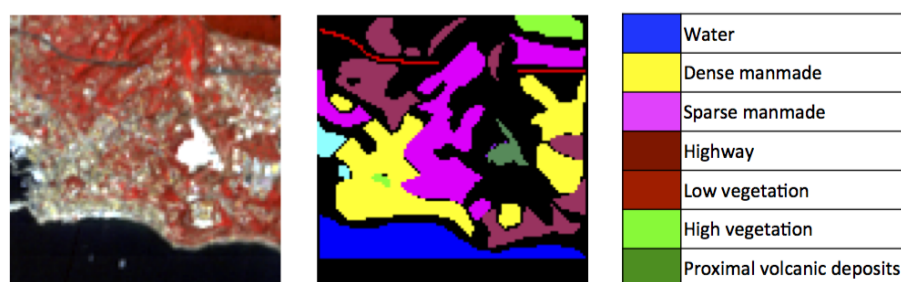
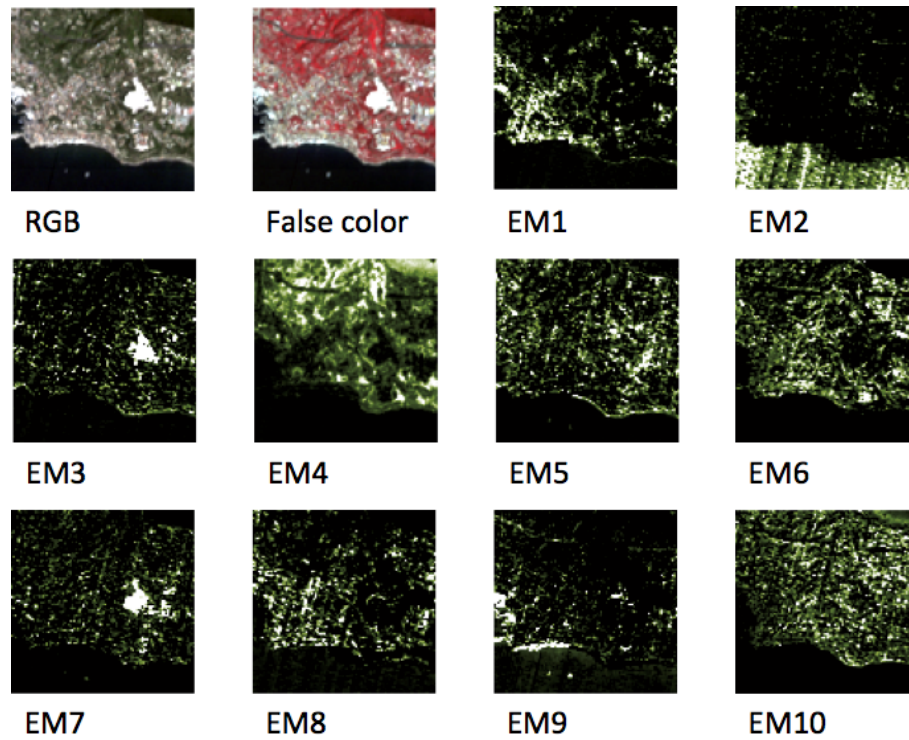
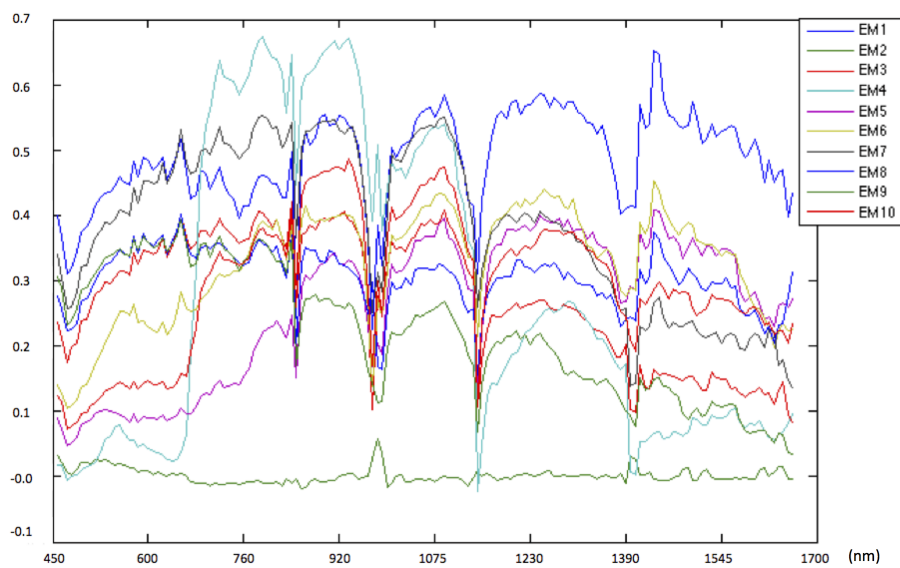


Figure 5. Hyperion image: ground truth.



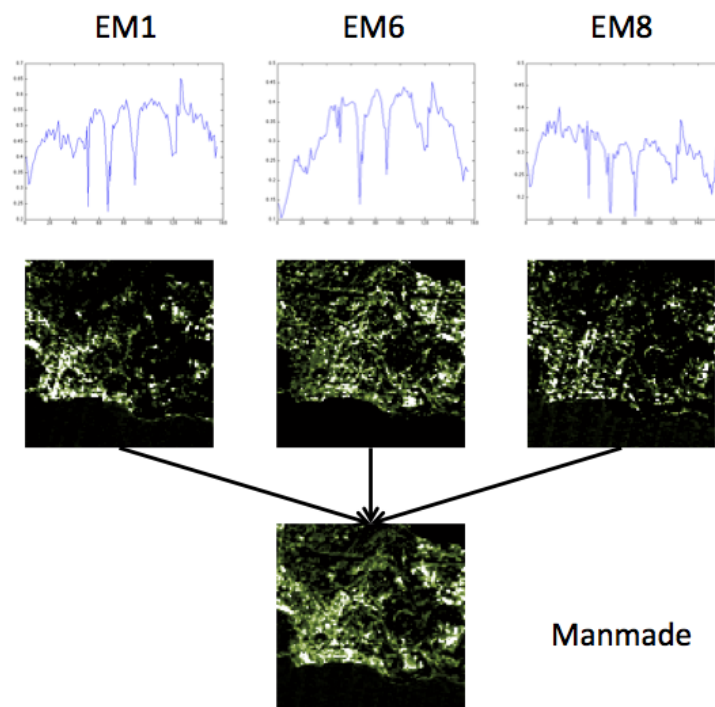
**Figure 6.** Hyperion image: fractional abundances obtained from the NLPCA derived data corresponding to the obtained endmembers (EM1, EM2,...,EM10).



**Figure 7.** Spectral signatures of the endmembers obtained with the proposed method for the Hyperion image.

From the obtained endmembers, it can be noted that EM1, EM6 and EM8 present similar spectral signatures, which, according to the available ground truth, represent different extents of man-made structures. Then, a new endmember representing man-made structures is obtained by summing up these three endmembers (Figure 8). In a similar way, both EM4 and EM10 could be associated to different types of vegetations. Finally, EM2 and EM5 clearly refer to water and bare soil, respectively. The remaining three endmembers EM3, EM7 and EM9 are the most interesting from a volcanic products deposit detection point of view. Proximal volcanic material is collectively identified by EM3

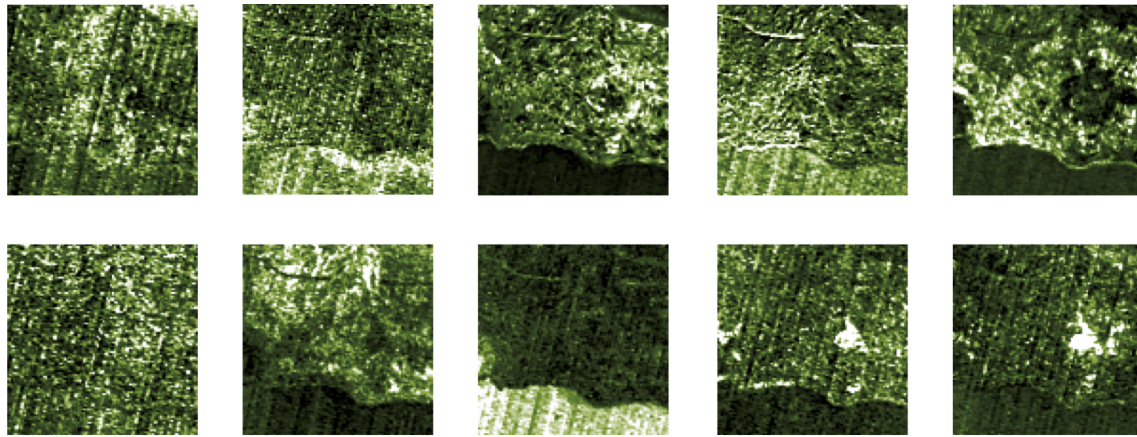
and 7, while distal material, in the coastline area and more northwestwards, by EM9. The spectral signatures of EM3, 7 and 9, while each one different with respect to the others, present a similar lower reflectance (higher absorption) in the range between about 950 nm and 1000 nm. This identifies similar material, while in morphological (surface structure, rugosity, macroscopic properties of the deposit) and, probably, detailed mineralogical composition. Further studies are ongoing, starting with this pilot analysis, to detect the precise composition of the structures identified with these endmembers, using their retrieved spectral signatures.



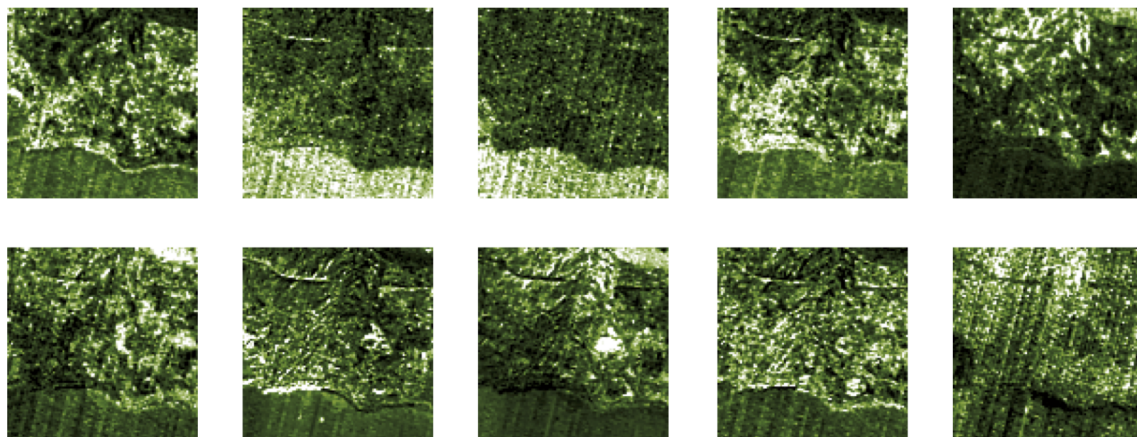
**Figure 8.** EM1, EM6 and EM8 combined to form a new endmember representing manmade surfaces.

For sake of comparison, the endmember extraction (N-FINDR) and abundance estimation (SUnSAL) phases have been applied also to the original image and the low-dimensionality image obtained through the use of linear PCA, respectively. In particular, for the PCA, there have been selected nine linear principal components in order to have a comparison with the NLPCA. The obtained abundance maps are reported in Figures 9 and 10. In a first analysis, it can be noted that, in both cases, the noise have a strong impact on the abundance maps. As for the abundance maps obtained with the original image, EM1, EM4 and EM7 can be associated to manmade surfaces. EM3 seems to be associated to water surfaces while EM2 can be associated to low reflective surfaces. EM5 and EM6 are related to vegetation while EM9 could be associated to bare soil. EM8 is associated to both volcanic deposits and also dense vegetation. Finally, EM10 is strongly influenced by noise and cannot be associated to any material. On a more deep analysis, it can be noted that volcanic products deposit cannot be identified in one endmember but are present in both EM7 and EM8. A similar result can be obtained analyzing the abundances derived from the nine linear PCs. However, the amount of noise affecting the abundance maps is higher. In both cases, the volcanic products deposit are not completely identified.

The main result of this first experiment is that, using an atmospheric correction routine, the volcanic plume features have been apparently removed and the characterisation of the surface volcanic material deposit is then possible as a lesser complex spectroscopic interpretation problem, and with a limited output space dimensionality.



**Figure 9.** Fractional abundances obtained from the original Hyperion image.



**Figure 10.** Hyperion image: fractional abundances obtained from the PCA derived image.

### 3.2. Kilauea Volcano

A second experiment has been carried out using a hyperspectral image obtained with the AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) sensor. AVIRIS is an airborne optical sensor delivering calibrated images of the upwelling spectral radiance in 224 contiguous spectral bands with wavelengths from 0.4 to 2.5  $\mu\text{m}$ . The image used in this case-study has been acquired in 2007 over the Puu Oo crater, in the eastern rift zone of the Kilauea volcano of the Hawaiian Islands Figure 11. The volcanic activity is usually characterized by emissions of both ash and sulphur dioxide. This image was taken during moderate activity of the volcano.

In this case, an atmospheric correction is not applied. By opposition with respect to the previous case, this processing chain allows to actually identify plume features by means of the endmember extraction.

As done for the previous experiment, the AVIRIS radiance image has been projected in the feature space by means of NLPCA. In this case, the optimal topology is found to have 224 nodes in the input/output layers, 110 nodes in the coding/decoding layers and 10 nodes in the bottleneck layer. Then, as for the CF case-study, also in this case, the dimensionality of the hyperspectral input vector is reduced by about the 95% by the nonlinear compression of the implemented AANN. Subsequently, seven endmembers and the corresponding fractional abundance maps have been identified through the use of the N-FINDR and SUNSAL algorithms (Figures 12 and 13).

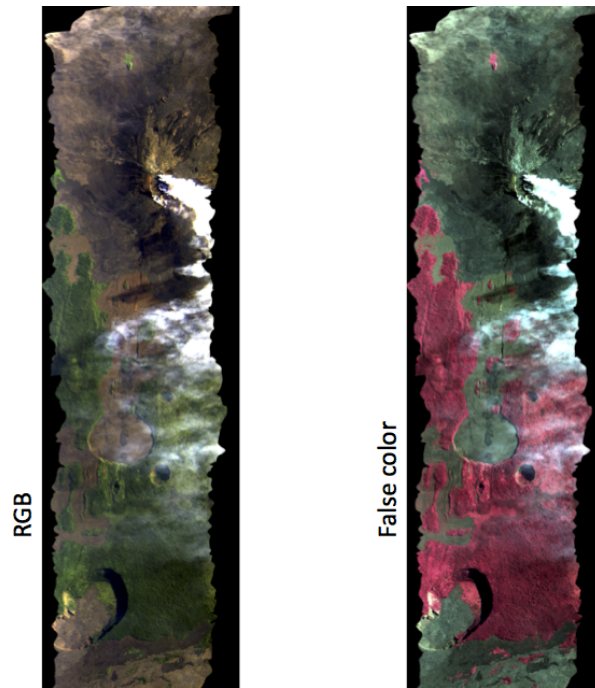


Figure 11. Color and false color representations of the AVIRIS image.

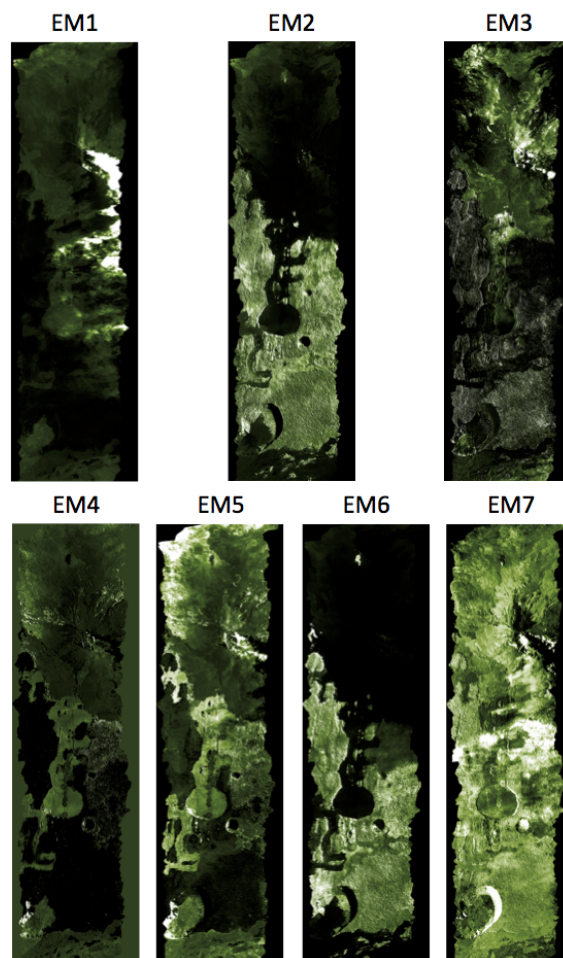
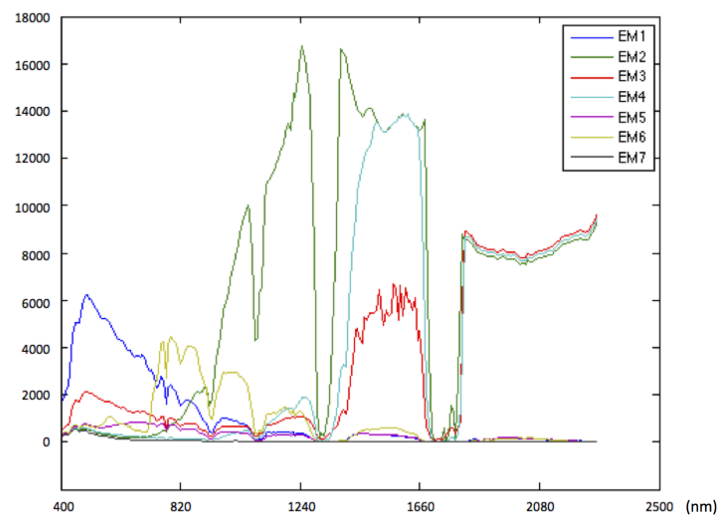


Figure 12. Abundance maps obtained from the AVIRIS image using the NLPCA approach.



**Figure 13.** Spectral signatures of the endmembers obtained with the proposed method for the AVIRIS image.

Surface features have been identified, as for the previous case. Thanks to comparison with ground truth, endmembers EM2 and 6 are associated with different types of vegetation and EM4 is associated to bare soil, i.e., basalt. Surface volcanic deposit are also identified, as for EM3 that is associated with tephra deposits. Differently with respect to the CF case, volcanic plume features here are also identified, as they are not screened out by the atmospheric correction carried out for the previous case. Endmember EM1 can be associated with water vapour emissions and the subsequent condensation to liquid clouds of volcanic origin. It tracks very well the plume geometry, as water vapour is a major volcanic effluent and is very apparent in the extracted image using only EM1. Looking at the spectral signature of this endmember, the rapid decrease of the reflectivity between the visible range (maximum at about 500 nm) and the near infrared, point at scattering processes by relatively small particles, as freshly nucleated water droplets following the emission by the volcano. Endmember EM5, while presenting several surface structures, including volcanic deposit, as for EM3, still holds a certain representativity of the plume object, in particular in the near range, as visible from Figure 12. The spectral signature of EM5 has a larger spectral structure than EM1, in the visible range, and so it can be argued that it carries information about larger ash particles. This hypothesis is supported, in conjunction with these spectroscopic/scattering considerations on the spectral signature of this endmember, by the fact that: (a) the plume identification of EM5 is confined to the near-crater area, thus suggesting coarse particles with quick sedimentation, and (b) the simultaneous identification of surface volcanic deposit features, probably with mineralogical composition consistent with the plume-ash discussed here. Further studies are required to attempt finer attribution of these mixed plume/surface endmembers' identification to specific volcanic products, even if these first results are encouraging towards the application of the methodology. Finally, EM7 is associated to the different illumination paths (shadows), and then carry very limited information on atmospheric or surface features.

Similarly to the previous experiment, a comparison with the results of the spectral unmixing (N-FINDR + SUnSAL), applied to the original and the PCA derived images, respectively, has been carried out. According to the abundance maps reported in Figures 14 and 15, it can be noted that, similarly to the NLPCA case, in both cases, there are endmembers related to different vegetation surfaces, bare soil and water vapour. However, the main difference relies in the way the volcanic materials are detected. In particular, with linear approaches, it is not possible to detect all the volcanic materials present in the scene, as it was possible with the NLPCA approach.

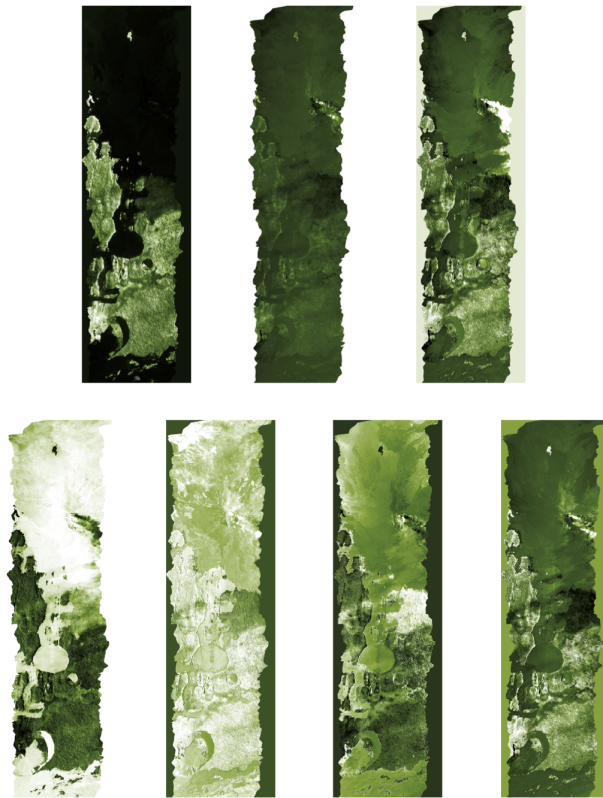


Figure 14. Abundance maps obtained from the original AVIRIS image.

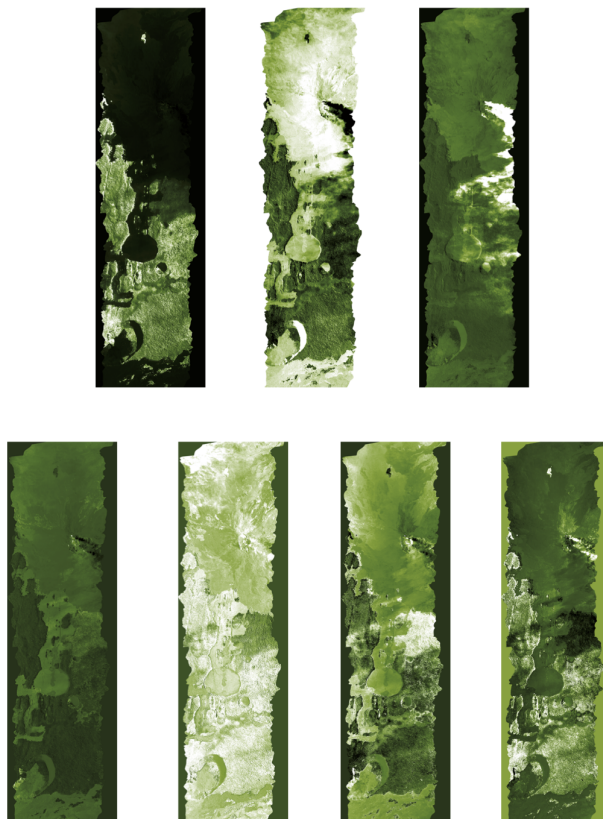


Figure 15. AVIRIS image: abundance maps obtained from seven linear PCs.

#### 4. Conclusions

In this paper, a nonlinear spectral unmixing procedure is tested, using AANN, to separate the various components of remote sensing imagery of volcanoes and their products. The problem of separating background, surface deposits of volcanic ejecta and airborne plumes of volcanic effluents, basing on their individual spectral signatures, from a given remote sensing image, is tackled. The method is tested on two different cases: (1) Campi Flegrei area (deposits and fumaroles) and (2) Kilauea volcano (deposits and plume for a moderate eruption). The dimensionality of the spectroscopic problem is reduced of up to 95%, using narrow-bottleneck AANNs, thus removing the redundancies of the hyperspectral input data. An atmospheric correction procedure is applied to the first case but not to the second. The Campi Flegrei case-study is very suitable to try the complete removal of the plume interference, due to the limited effluents concentration for this case. The atmospheric correction has enabled the access to the identification of volcanic and background features for this case. On the contrary, the thicker plume for the Kilauea case-study made this case very well adapted to investigate both plume and surface volcanic products. Then, an atmospheric correction is not performed to retain the airborne plume information for the subsequent analyses. For Kilauea, the method allows the identification partial separation of different plume (e.g., water vapour/condensed droplets, ash) and surface tephra deposits. While more analyses and validation efforts, using ground truth or complementary remote observations, are needed, this pilot study demonstrate the potential of this methodology to tackle different unmixing problems linked to the identification of volcanic products from remote-sensing imagery. This is also confirmed by comparing the results obtained with the proposed method with those obtained using classical linear approaches. However, it is important to state that the obtained endmembers are extreme points only in the feature space but they might not be in the original observation space. For this reason, the obtained abundances may not lead to objective quantification of the endmembers. Finally, if not correctly designed, the dimensionality reduction using the NLPCA approach, may lead to endmembers that have physical meaning but do not exist in the real world. This is the case of features representing the illumination or in some cases sensor artifacts.

**Acknowledgments:** The work of Pasquale Sellitto has been partially funded by the FP7 Med-SuV and the INGV SMED projects. The work of Giorgio Licciardi and Jocelyn Chanussot has been funded by the ANR ASTRID (project APHYPIS) under grant ANR-16-ASTR-0027-01.

**Author Contributions:** G.A.L. conceived the experiments; G.A.L., A.P., P.S. and J.C. analysed the data and discussed the results; G.A.L. and P.S. wrote the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Von Glasow, R. Atmospheric chemistry in volcanic plumes. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 6594–6599.
2. Prata, A.J. Infrared radiative transfer calculations for volcanic ash clouds. *Geophys. Res. Lett.* **1989**, *16*, 1293–1296.
3. Gangale, G.; Prata, A.J.; Clarisse, L. On the infrared spectral signature of volcanic ash. *Remote Sens. Environ.* **2010**, *114*, 414–425.
4. Karagulian, F.; Clarisse, L.; Clerbaux, C.; Prata, A.J.; Hurtmans, D.; Coheur, P.F. Detection of volcanic SO<sub>2</sub>, ash, and H<sub>2</sub>SO<sub>4</sub> using the Infrared Atmospheric Sounding Interferometer (IASI). *J. Geophys. Res. Atmos.* **2010**, *115*, 1–10.
5. Clarisse, L.; Hurtmans, D.; Clerbaux, C.; Hadji-Lazaro, J.; Ngadi, Y.; Coheur, P.F. Retrieval of sulphur dioxide from the infrared atmospheric sounding interferometer (IASI). *Atmos. Meas. Tech.* **2012**, *5*, 581–594.
6. Carboni, E.; Grainger, R.; Walker, J.; Dudhia, A.; Siddans, R. A new scheme for sulphur dioxide retrieval from IASI measurements: Application to the Eyjafjallajökull eruption of April and May 2010. *Atmos. Chem. Phys.* **2012**, *12*, 11417–11434.
7. Li, C.; Krotkov, N.A.; Carn, S.; Zhang, Y.; Spurr, R.J.D.; Joiner, J. New-generation NASA Aura Ozone Monitoring Instrument (OMI) volcanic SO<sub>2</sub> dataset: Algorithm description, initial results, and continuation with the Suomi-NPP Ozone Mapping and Profiler Suite (OMPS). *Atmos. Meas. Tech.* **2017**, *10*, 445–458.



8. Sellitto, P.; Legras, B. Sensitivity of thermal infrared nadir instruments to the chemical and microphysical properties of UTLS secondary sulfate aerosols. *Atmos. Meas. Tech.* **2016**, *9*, 115–132.
9. Sellitto, P.; Sèze, G.; Legras, B. Secondary sulphate aerosols and cirrus clouds detection with SEVIRI during Nabro volcano eruption. In Proceedings of the EGU General Assembly 2016, Vienna, Austria, 17–22 April 2016. (In typesetting for International Journal of Remote Sensing)
10. Prata, A.J. Observations of volcanic ash clouds in the 10–12  $\mu\text{m}$  window using AVHRR/2 data. *Int. J. Remote Sens.* **1989**, *10*, 751–761.
11. Prata, A.J.; Kerkmann, J. Simultaneous retrieval of volcanic ash and  $\text{SO}_2$  using MSG-SEVIRI measurements. *Geophys. Res. Lett.* **2007**, *34*, L05813.
12. Corradini, S.; Merucci, L.; Prata, A.J.; Piscini, A. Volcanic ash and  $\text{SO}_2$  in the 2008 Kasatochi eruption: Retrievals comparison from different IR satellite sensors. *J. Geophys. Res. Atmos.* **2010**, *115*, doi:10.1029/2009JD013634.
13. Champion, R.; Salerno, G.G.; Coheur, P.F.; Hurtmans, D.; Clarisse, L.; Kazahaya, K.; Burton, M.; Caltabiano, T.; Clerbaux, C.; Bernard, A. Measuring volcanic degassing of  $\text{SO}_2$  in the lower troposphere with ASTER band ratios. *J. Volcanol. Geotherm. Res.* **2010**, *194*, 42–54.
14. Rix, M.; Valks, P.; Hao, N.; van Geffen, J.; Clerbaux, C.; Clarisse, L.; Coheur, P.F.; Erbertseder, T.; Zimmer, W.; Emmadi, S. Satellite Monitoring of Volcanic Sulfur Dioxide Emissions for Early Warning of Volcanic Hazards. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2009**, *2*, 196–206.
15. Picchiani, M.; Chini, M.; Corradini, S.; Merucci, L.; Sellitto, P.; Del Frate, F.; Stramondo, S. Volcanic ash detection and retrievals using MODIS data by means of neural networks. *Atmos. Meas. Tech.* **2011**, *4*, 2619–2631.
16. Piscini, A.; Carboni, E.; Del Frate, F.; Grainger, R.G. Identifying volcanic endmembers in hyperspectral images using spectral unmixing. In Proceedings of the SPIE 9242, Remote Sensing of Clouds and the Atmosphere XIX, and Optics in Atmospheric Propagation and Adaptive Systems XVII, Amsterdam, The Netherlands, 22 September 2014.
17. Piscini, A.; Picchiani, M.; Chini, M.; Corradini, S.; Merucci, L.; Del Frate, F.; Stramondo, S. A neural network approach for the simultaneous retrieval of volcanic ash parameters and  $\text{SO}_2$  using MODIS data. *Atmos. Meas. Tech.* **2014**, *7*, 4023–4047.
18. Piscini, A.; Carboni, E.; Del Frate, F.; Grainger, R.G. Simultaneous retrieval of volcanic sulphur dioxide and plume height from hyperspectral data using artificial neural networks. *Geophys. J. Int.* **2014**, *198*, 697.
19. Keshava, N. A Survey of Spectral Unmixing Algorithms. *MIT Linc. Lab. J.* **2003**, *14*, 55–78.
20. Keshava, N.; Mustard, J.F. Spectral Unmixing. *IEEE Signal Process. Mag.* **2002**, *19*, 44–57.
21. Singer, R.B.; McCord, T.B. Mars: Large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance. In Proceedings of the Lunar and Planetary Science Conference, Houston, TX, USA, 19–23 March 1979; pp. 1835–1848.
22. Hapke, B. Bidirection reflectance spectroscopy. I. Theory. *J. Geophys. Res.* **1981**, *86*, 3039–3054.
23. Clark, R.N.; Roush, T.L. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *J. Geophys. Res.* **1984**, *89*, 6329–6340.
24. Fan, W.; Hu, B.; Miller, J.; Li, M. Comparative study between a new nonlinear model and common linear model for analyzing laboratory simulated-forest hyperspectral data. *Int. J. Remote Sens.* **2009**, *30*, 2951–2962.
25. Altmann, Y.; Dobigeon, N.; Tourneret, J.Y. Bilinear models for nonlinear unmixing of hyperspectral images. In Proceedings of the 2011 3rd Workshop on Hyperspectral Image Signal Processing: Evolution in Remote Sensing (WHISPERS), Lisbon, Portugal, 6–9 June 2011.
26. Kulbelka, P.; Munk, F. Reflection characteristics of paints. *Z. Tech. Phys.* **1931**, *12*, 593–601.
27. Shkuratov, Y.; Starukhina, L.; Hoffmann, H.; Arnold, G. A model of spectral albedo of particulate surfaces: Implications for optical properties of the Moon. *Icarus* **1999**, *137*, 235–246.
28. Plaza, J.; Plaza, A. Spectral mixture analysis of hyperspectral scenes using intelligently selected training samples. *IEEE Geosci. Remote Sens. Lett.* **2010**, *7*, 371–375.
29. Altmann, Y.; Dobigeon, N.; McLaughlin, S.; Tourneret, J.Y. Nonlinear unmixing of hyperspectral images using radial basis functions and orthogonal least squares. In Proceedings of the IEEE International Conference on Geoscience and Remote Sensing (IGARSS), Vancouver, BC, Canada, 24–29 July 2011; pp. 1151–1154.
30. Licciardi, G.; Frate, F.D. Pixel Unmixing in Hyperspectral Data by Means of Neural Networks. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4163–4172.

31. Cybenko, G. Approximations by superpositions of sigmoidal functions. *Math. Control Signals Syst.* **1989**, *2*, 303–314.
32. Kramer, M.A. Nonlinear principal component analysis using autoassociative neural networks. *AIChE J.* **1991**, *37*, 233–243.
33. Sellitto, P. Artificial neural networks for spectral sensitivity analysis to optimize inversion algorithms for satellite-based earth observation: Sulfate aerosol observations with high-resolution thermal infrared sounders. In *Sensitivity Analysis in Earth Observation Modelling*; Petropoulos, G.P., Srivastava, P.K., Eds.; Elsevier: Amsterdam, The Netherlands, 2017; pp. 161–175.
34. Bioucas-Dias, J.M.; Plaza, A.; Dobigeon, N.; Parente, M.; Du, Q.; Gader, P.; Chanussot, J. Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 354–379.
35. Plaza, A.; Chang, C.I. An improved N-FINDR algorithm in implementation. In Proceedings of the SPIE 5806, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI, Orlando, FL, USA, 28 March 2005.
36. Bioucas-Dias, J.; Figueiredo, M. Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing. In Proceedings of the 2010 2nd Workshop on Hyperspectral Image Signal Processing: Evolution in Remote Sensing (WHISPERS), Reykjavik, Iceland, 14–16 June 2010.
37. Bishop, C. *Neural Networks for Pattern Recognition*; Oxford University Press: London, UK, 1995.
38. Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Netw.* **1991**, *4*, 251–257.
39. Licciardi, G.; Marpu, P.R.; Chanussot, J.; Benediktsson, J.A. Linear Versus Nonlinear PCA for the Classification of Hyperspectral Data Based on the Extended Morphological Profiles. *IEEE Geosci. Remote Sens. Lett.* **2011**, *9*, 447–451.
40. Datt, B.; Vicar, T.R.M.; Niel, T.G.V.; Jupp, D.L.B.; Pearlman, J.S. Preprocessing EO-1 Hyperion Hyperspectral Data to Support the Application of Agricultural Indexes. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1246–1259.
41. Lima, A.; Vivo, B.D.; Spera, F.J.; Bodnar, R.J.; Milia, A.; Nunziata, C.; Belkin, H.E.; Cannatelli, C. Thermodynamic model for uplift and deflation episodes (bradyseism) associated with magmatic-hydrothermal activity at the Campi Flegrei (Italy). *Earth Sci. Rev.* **2009**, *97*, 44–58.
42. Deino, A.L.; Orsi, G.; de Vita, S.; Piochi, M. The age of the Neapolitan Yellow Tuff caldera-forming eruption (Campi Flegrei caldera—Italy) assessed by <sup>40</sup>Ar/<sup>39</sup>Ar dating method. *J. Volcanol. Geotherm. Res.* **2004**, *133*, 157–170.
43. Orsi, G.; Di Vito, M.A.; Isaia, R. Volcanic hazard assessment at the restless Campi Flegrei caldera. *Bull. Volcanol.* **2004**, *66*, 514–530.
44. Chiodini, G.; Caliro, S.; Cardellini, C.; Granieri, D.; Avino, R.; Baldini, A.; Donnini, M.; Minopoli, C. Long-term variations of the Campi Flegrei, Italy, volcanic system as revealed by the monitoring of hydrothermal activity. *J. Geophys. Res. Solid Earth* **2010**, *115*, doi:10.1029/2008JB006258



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).