Physical Retrieval of Land Surface Emissivity Spectra from Hyper-Spectral Infrared Observations and Validation with In Situ Measurements

Guido Masiello 1,*, Carmine Serio 1,†, Sara Venafra 1, Giuliano Liuzzi 2, Laurent Poutier 3,§ and Frank-M. Göttsche 4,∥

1 School of Engineering, University of Basilicata, Viale Ateneo Lucano, 10, 85100 Potenza, Italy; carmine.serio@unibas.it (C.S.); sara.venafra@unibas.it (S.V.)
2 NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA; giuliano.liuzzi@nasa.gov
3 ONERA, The French Aerospace Lab, 2 avenue Edouard Belin, FR-31055 Toulouse, France; laurent.poutier@onera.fr
4 IMK-ASF, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany; frank.goettsche@kit.edu
* Correspondence: guido.masiello@unibas.it; Tel.: +39-0971-205158

Received: 17 May 2018; Accepted: 14 June 2018; Published: 20 June 2018

Abstract: A fully physical retrieval scheme for land surface emissivity spectra is presented, which applies to high spectral resolution infrared observations from satellite sensors. The surface emissivity spectrum is represented with a suitably truncated Principal Component Analysis (PCA) transform and PCA scores are simultaneously retrieved with surface temperature and atmospheric parameters. The retrieval methodology has been developed within the general framework of Optimal Estimation and, in this context, is the first physical scheme based on a PCA representation of the emissivity spectrum. The scheme has been applied to IASI (Infrared Atmospheric Sounder Interferometer) and the retrieved emissivities have been validated with in situ observations acquired during a field experiment carried out in 2017 at Gobabeb (Namib desert) validation station. It has been found that the retrieved emissivity spectra are independent of background information and in good agreement with in situ observations.

Keywords: emissivity spectrum; infrared; satellite; optimal estimation validation

1. Introduction

The problem of retrieving surface emissivity from infrared spectral observations is mainly that of separating it from temperature in the surface radiation emission: this is is commonly referred to as $T_s - \epsilon$ (Temperature-Emissivity) separation (e.g., see [1,2] and references therein).

Within the framework of high spectral resolution infrared observations from air planes and satellites, the problem has been addressed by [3,4], who arrived at a closed form, in which emissivity is separated from surface temperature and atmospheric emission by using the radiation reflected at the surface, which under a Lambertian model depends on emissivity alone. Thus, emissivity could be retrieved provided that the state of the surface-atmosphere was known. Assuming a suitable dependence of emissivity on wave number, the scheme can then be used to develop a least square procedure in which $T_s - \epsilon$ are simultaneously retrieved.

Similar approaches have been considered in previous studies from aircraft, ship and in situ observations of spectral radiance, usually acquired with Fourier Transform Spectrometers [5–8]. All of these methods inevitably assume that the atmospheric state is known from independent sources,
whereas the methodology we describe and demonstrate, simultaneously retrieves surface and atmospheric parameters from spectral information contained in the earth emission spectrum.

Several algorithms have been developed to derive emissivity in the infrared from satellite measurements [9–18]. The retrieval is generally accomplished with multi-stage algorithms, that separate temperature and emissivity using additional background emissivity information derived from laboratory measurements.

Another important issue concerning emissivity retrieval is the fact that the emissivity vector could potentially have a dimensionality as large as that of the observed radiance vector. For a given spectral observation, this would lead to an inverse problem with more unknowns than data points. Thus, the problem of effectively retrieving emissivity from infrared observations is fundamentally one of emissivity-dimension reduction, to which one possible solution was given by [19–21], through decomposition of the emissivity function on a suitable orthogonal basis, a technique dating back to [22].

Along these lines, a full physical retrieval algorithm has been developed by [23] who used a Fourier transform representation of the emissivity spectrum and the Fourier coefficients were retrieved simultaneously with surface temperature and atmospheric parameters.

The methodology developed in this study is based on [23], but with two important improvements: (1) the entire observed spectrum is used, (2) the Fourier analysis is changed to a Principal Component Analysis (PCA), which is much more parsimonious in terms of coefficients or scores while still properly representing the emissivity spectrum. The scheme will be exemplified for IASI (Infrared Atmospheric Sounder Interferometer) [24], although the methodology can be validly applied to any hyper-spectral instrumentation.

It should be stressed that the emissivity seen by satellites depends on the geometry of the scene as imaged by the instrument optics. The scene can contain different elements, such as soil, bedrock and vegetation. In addition, vegetation may be in different phenological states (dry, senescent, green) and the surface may be altered by seasonal snow cover. This also limits a full validation of the emissivity retrievals with in situ observations, since the area observed by the ground instrumentation is not necessarily representative of the area observed by the satellite and the viewing geometries associated with the two sensors are rarely identical. For this reason, the validation study presented here was performed over highly homogeneous and flat gravel plains around the permanent Land Surface Temperature (LST) validation station ‘Gobabeb’ in the Namib desert [25,26] operated by Karlsruhe Institute of Technology (KIT) within the framework of EUMETSAT’s Satellite Application Facility on Land Surface Analysis (LSA SAF).

2. Data and Methods

IASI emissivity retrievals were validated with in situ emissivity spectra acquired during an international Land Surface Temperature (LST) field inter-comparison experiment (FICE), which was organised within the framework of ESA’s Fiducial Reference Measurements for validation of Surface Temperature from Satellites (FRM4STS) project [27] and took place between the 17th and 24th of June 2017 at Gobabeb, Namibia.

2.1. Validation Site

Gobabeb Research and Training Centre (GRTC; www.gobabebtrc.org, accessed on 19 June 2018) in Namibia is the only permanently staffed desert research station worldwide. GRTC is located on a sharp transition between the vast Namib sand sea with its up to 300 m high dunes and adjacent gravel plains: this natural boundary is maintained by irregular flows of the ephemeral Kuiseb River (a few days every other year), which wash the advancing sand into the South Atlantic Ocean. Due to the hyper-arid desert climate [28], the site is spatially and temporally highly stable and, therefore, ideal for long-term validation studies of satellite products [16]. The long-term average annual temperature at Gobabeb (GBK) is 21.1 °C [29] whereas the average annual precipitation is less than 100 mm [30] and highly variable [31]. Consequently, the relatively frequent fog events are of special importance for
the water balance of the Namib [30]. Continuous in-situ measurements are performed at Karlsruhe Institute of Technology (KIT) two permanent LST validation stations ‘GBB Wind’ (23.551°S, 15.051°E, 450 m asl) and ‘GBB Plains’ (23.519°S, 15.083°E, 450 m asl). GBB Wind (see Figure 1) uses a 30 m high wind profiling tower about 2 km north-east of GRTC, while GBB Plains uses a 25 m high telescopic mast about 7 km north-east of GRTC.

![Figure 1. KIT’s permanent validation station GBB Wind on the 16 June 2017.](image)

Both stations are at the edge of several thousand km$^2$ of gravel plains, which are covered by a highly homogeneous mixture of gravel, sand and sparse desiccated grass. Nevertheless, for reliable product validation the effect of the small scale variation of surface materials (e.g., dry grass, rock outcrops) and topography needs to be fully characterized. Using a mobile radiometer system, several field experiments were performed during which a radiometer was driven along tracks of up to 40 km length across the gravel plains. The results showed a high level of homogeneity and a stable relationship between GBB Wind LST and the LST along the tracks with biases between $-0.1$ and $0.8$ °C [25].

Clear sky conditions are preferable for field measurements since down-welling radiance is then easier to determine and varies relatively slowly and smoothly. Furthermore, LST retrieval from passive Thermal Infrared (TIR) satellite sensors also requires clear-sky situations. Since Gobabeb is located in the Namib Desert it offers frequent clear sky conditions almost all year around, which is ideal for LST validation.

The in situ emissivity observations used here for validation purposes were obtained on the Namib gravel plains near Gobabeb LST validation station (Figure 2) during an international LST field inter-comparison experiment. Before the experiment the Namib experienced several years without significant rainfall and the gravel plains were exclusively covered by gravel and sand, i.e., only a minimum amount of dry grass was present (also see Figure 1). The various measurements performed during the field experiment are described in some detail in [27].
Figure 2. Location of in situ measurements and collocated IASI soundings. The tags from A to D show the location of in situ measurements on 17, 22, 23 and 24 June 2017, respectively. Red circles correspond to IASI footprints over desert, those green over the gravel plain; in blue the IASI soundings best collocated with in situ observations.

2.2. Data

The in-situ emissivity spectra were obtained from measurements with a Fourier Transform Infrared (FTIR) spectrometer by ONERA, the French Aerospace Lab; the emissivity spectra cover the range from 750 to 1250 cm$^{-1}$ (8–12 µm) with a spectral resolution of 4 cm$^{-1}$.

The applied methodology had been previously evaluated in the 3–13 µm (≈ 700–3000 cm$^{-1}$) spectral range by comparing in-lab reference measurements to outdoor emissivity retrievals for a large set of surfaces including various mineral soils [32]. It led to an overall rms emissivity uncertainty in the 8–12 µm domain better than 0.02. The technique consists in measuring both the ground upwelling radiance and the downwelling irradiance. The ill-conditioned emissivity-temperature separation is then solved using a spectral smoothness constraint [33] exploiting the higher variability of the spectrum of atmospheric irradiance compared to emissivity spectra of natural surfaces.

The spectra are acquired by a BOMEM MR304SC FTIR spectroradiometer (Québec City, Canada) equipped with a 75 mrad Field of View telescope and a 45° flat mirror. With this setup, the ground target surface is viewed at nadir and the diameter of the analyzed area is approximately 15 cm. The downwelling irradiance is measured using a Labsphere Infragold 25 × 25 cm standard reflector (Sutton, NH, USA) at the target’s place (see Figure 3), so that the instrumentation’s contribution to the downwelling irradiance is properly compensated. The reflector self emission is also corrected for using its temperature monitored by a PRT probe (Platinum Resistance Thermometers) associated with its spectral reflectance measured in the lab. The two acquisitions on the target sample and the reflector are done sequentially within typically less than 1 min at a 4 cm$^{-1}$ spectral resolution. The temporal variations of the atmospheric conditions are assumed negligible in this time interval. The radiometric calibration uses two acquisitions of the same MIKRON M345 10 × 10 cm blackbody set (Santa Clara, CA, USA) at two different temperatures, done alternatively before and after the measurements. For these measurements, the flat mirror is tilted and the blackbody active surface is vertical. The blackbody apparent emissivity is measured in the lab and its reflective contribution is supposed to come from an environment at a brightness temperature equal to the ambient temperature, downloaded from the Gobabeb meteorological station (http://www.sasscalweathernet).
This calibration process has been validated during a comparison of blackbodies and radiometers completed at National Physical Laboratory (NPL) in June 2016 [34,35]. 45 measurements have been acquired on the gravel plain, at different locations around the station in order to capture as much of the spatial emissivity variability as possible. These in situ observations have been merged into 10 different data sets. Seven data sets were acquired on the 17 June 2017 in the vicinity of the GBB wind mast (spot A in Figure 2), and the remaining three on the 22–24 June 2017 correspond to the GBB plains mast, GRTC and along the track Gobabeb—Mirabib respectively (namely spots B, C and D in Figure 2). The mean and standard deviation of each data set were computed per wavelength; standard deviation provides a measure of variability within a set combined with the measurement uncertainty. For the first 7 sets acquired on the 17 June, the average emissivity spectra for the ten sets and their associated variabilities (standard deviation) are shown in Figure 4. It should be noted that for the sets 1 to 7 (taken on 17 June 2017) the uncertainty is almost constant along the spectral coverage and nearly equal to 0.015. Conversely, the measurements on the remaining three days show a much larger variability, which can reach 0.03\% or 3\% in the core of the quartz reststrahlen band at 8.6 \( \mu \text{m} \) (1150 cm\(^{-1}\)), due to the spatial variability within each sampling. The averaged emissivity has been used for the comparison with IASI retrievals.

![Figure 3. Instrumentation setting used for in situ measurements of emissivity soundings.](image)

![Figure 4. (a) Mean emissivity of the ten sets of in situ measurements. (b) Standard deviation.](image)

IASI soundings have been collected over the days 17 to 24 June 2017 for a target area, which overlaps the period of the in situ measurements. The IASI instrument has a spectral coverage extending from 645 to 2760 cm\(^{-1}\), with a sampling interval \( \Delta \sigma = 0.25 \text{ cm}^{-1} \) providing 8461 data...
points or channels for each single spectrum. IASI is a cross-track scanner, with thirty effective field
of regard (FOR) per scan, which span an angle range of \( \pm 48.33^\circ \) on either side of nadir; the two
symmetric nadir FORs are collected at angles of \( \pm 1.67^\circ \). Each FOR consists of a \( 2 \times 2 \) matrix of so-called
instantaneous fields of view (IFOV). Each IFOV has a diameter of 14.65 mrad, which corresponds to a
ground resolution of 12 km at nadir for a satellite altitude of 819 km. The \( 2 \times 2 \) matrix is centred on
the viewing direction. At nadir the FOR of 4 IASI pixels project to a ground a square area of \( \approx 50 \times 50 \text{ km}^2 \).
More details on IASI and its mission objectives can be found in [24].

IASI footprints corresponding to desert sand (14 soundings) and the gravel plain (20 soundings)
are shown in Figure 2. Those on the desert sand were used to check the sensitivity of the
retrieval algorithm to different emissivity features. The IASI soundings correspond to view angles
between 5 and 55\(^\circ\). This should be kept in mind when comparing to \textit{in situ} observations, which are
taken at nadir. Furthermore, the IASI footprint approximates a circle with a diameter of 12 km at nadir,
whereas the \textit{in situ} measurements represent point-like observations, which were collected over regions
with an extend of about 6 km.

Six IASI soundings (blue circles in Figure 2) were used in comparison against the \textit{in situ}
measurements, i.e., those best collocated with the ground-based observations. Only clear sky IASI
measurements were considered. IASI clear sky soundings were identified through the AVHRR-based
cloud mask, which is provided by EUMETCast along with IASI L1C radiances. This cloud mask was
further improved by using the IASI stand alone cloud detection algorithm developed by [36,37]).

2.3. Retrieval Methodology

The retrieval methodology is based on what we call the \( \phi \)-IASI package, which is the combination
of a forward mode, which can compute both upwelling and downwelling spectral radiances [38–42] and
an inverse methodology [43] based on Optimal Estimation [44]. The approach has been documented
and validated in a series of papers [45–49].

Applications of the methodology to surface emissivity have been explicitly dealt with [23,50,51].
In its most updated version (see e.g., [52–54]), the full IASI spectrum (8461 channels) is used to
simultaneously retrieve, the surface temperature \( T_s \), emissivity spectrum \( \varepsilon \), the atmospheric
profiles of temperature \( T \), water vapour \( Q \), ozone \( O \), HDO, and CO\(_2\), and the column amount
of N\(_2\)O, CO, CH\(_4\), SO\(_2\), HNO\(_3\), NH\(_3\), OCS and CF\(_4\).

2.3.1. The Forward Model

The forward model we use in this paper [38] computes the upwelling spectral radiance \( R(\sigma) \)
(with \( \sigma \) the wave number) at the monochromatic level. The monochromatic spectrum is convolved
with the spectral response function of the given instrument (e.g., IASI) to obtain the appropriate spectral
radiance. For the case of IASI, the spectral radiance can be cast in a vector \( R \) of \( M = 8461 \) elements,
as many as the IASI channels.

The radiative transfer embedded in \( \sigma \)-IASI takes the form,

\[
R(\sigma) = R_s(\sigma) + R_a(\sigma) + R_r(\sigma)
\]

(1)

where \( R \) is the upwelling radiance decomposed in its surface term at the top-of-atmosphere, \( R_s \),
amospheric component, \( R_a \) and surface reflected part, \( R_r \), respectively. All quantities depend on
the wave number \( \sigma \) and the dependence over the directional angle is implicit. In its latest version,
\( \sigma \)-IASI can deal with Lambertian and/or specular surfaces and the solar radiation is properly taken
into account (e.g., [55]). The explicit inclusion of the term \( R_r(\sigma) \) is important for emissivity retrieval,
because it depends on emissivity alone, while the surface term, \( R_s(\sigma) \) depends on the product of
emissivity and the Planck function computed at the surface temperature.

The derivation and the explicit form of \( R_s(\sigma), R_a(\sigma), R_r(\sigma) \) we deal with in \( \sigma \)-IASI can be found
in [23] and they are not repeated here for the sake of brevity. For the right understanding of the present
paper it suffices to say that the forward model is a suitable non linear function \( F \), which relates the upwelling radiance vector, \( \mathbf{R} \) to the surface-atmosphere state vector, \( \mathbf{v} \). In a vector-matrix notation we have

\[
\mathbf{R} = F(\mathbf{v})
\]

where the radiance vector is made up with \( M \) radiance elements \( R(\sigma_i) \) computed at \( M \) wave numbers \( \sigma_i, \ i = 1, \ldots, M \). As said, for IASI we have \( M = 8461 \).

2.3.2. The Inverse Approach

The inverse methodology module is \( \delta \)-IASI \([43,52,54]\) and it yields an estimate \( \hat{\mathbf{v}} \) for a given set of observed radiances, \( \mathbf{R} \). Until 2016, the emissivity spectrum over the IASI range was parameterized through a Fourier Transform, whereas in the most up-to-date \( \delta \)-IASI version we use a truncated Principal Component Analysis (PCA) transform. The first twenty PCA scores are retained, hence retrieved. The use of PCA instead of Fourier is not a matter of retrieval accuracy, when rather a problem of compact representation of the emissivity spectrum. The Fourier series tends to converge slowly in comparison to the PCA transform. This effect is exemplified in Figure 5, where we show the emissivity retrieval for a set of 70 IASI soundings over the Algerian, Sahara desert in June 2007. It is seen that the Fourier expansion, to extract the same information as that corresponding to 20 PC scores, needs 450 coefficients (a factor more than 20!). For this reason, the Fourier approach has been dismissed in favour of PCA.

![Figure 5](image_url)

Figure 5. (a) Exemplifying the emissivity retrieval with Fourier series and PCA expansion. The results have been averaged over 70 IASI soundings recorded over a target area in the Algerian, Sahara desert in June 2007. Note how 20 PC scores are as much effective as 450 Fourier coefficients. (b) Difference between PCA and Fourier.

In its mathematical aspects, the PCA retrieval strategy for emissivity mimics the previous Fourier approach \([23]\). However, the mathematical details are discussed and presented below in some detail for the sake of clarity.

To retrieve the emissivity spectrum we have to properly define and compute a linearized form of the forward model around a suitable First Guess state vector, \( \mathbf{v}_o \). We have,

\[
\mathbf{R} - \mathbf{R}_o = \mathbf{K} (\mathbf{v} - \mathbf{v}_o)
\]
with \( \mathbf{R}_0 = F(\mathbf{v}_0) \). The size of the full state vector will be denoted by \( N \), therefore the Jacobian matrix \( \mathbf{K} \) has size \( M \times N \) and is computed according to

\[
\mathbf{K} = \frac{\partial F}{\partial \mathbf{v}} |_{\mathbf{v} = \mathbf{v}_0} \quad (4)
\]

Because the derivative is additive, the part related to the emissivity in the state vector and that to the other parameters can be separated,

\[
\mathbf{K} = (\mathbf{K}_1, \mathbf{K}_2) \quad (5)
\]

and

\[
\mathbf{v} - \mathbf{v}_0 = \begin{pmatrix} \mathbf{v}_1 - \mathbf{v}_{o1} \\ \mathbf{v}_2 - \mathbf{v}_{o2} \end{pmatrix} \quad (6)
\]

where the subscripts 1 and 2 refer to the atmospheric and emissivity components, respectively. The number of atmospheric parameters will be denoted by \( N_1 \), while that corresponding to emissivity by \( N_2 \). Of course, \( N_1 + N_2 = N \). For convenience the atmospheric parameters also include the surface temperature.

The Jacobian \( \mathbf{K}_1 \) is computed as usual as the derivative of the radiance, \( R(\sigma) \), with respect to the surface temperature and atmospheric parameters. In the same way, \( \mathbf{K}_2 \) is computed by differentiating \( R(\sigma) \) with respect to emissivity. However, the computation of \( \mathbf{K}_2 \) is less straightforward because we first transform emissivity to project it into an unbounded space (spanning from \(-\infty\) to \(+\infty\)) and then reduce its dimensionality.

We achieve this by transforming emissivity with the logit function, \( z(i) = \text{logit}(\epsilon(i)) = \log \left( \frac{\epsilon(i)}{1 - \epsilon(i)} \right) \), \( i = 1, \ldots, M \) \( (7) \), whose inverse is,

\[
\epsilon(i) = \frac{\exp(z(i))}{1 + \exp(z(i))}, \quad i = 1, \ldots, M \quad (8)
\]

where, to simplify the notation, we have written \( \epsilon(i) \) instead of \( \epsilon(\sigma_i) \) (with \( \sigma_i \) the wave number), and where \( M \) is the number of radiance data points or channels. We stress that at this stage the emissivity spectrum is represented with a vector of size \( M \), i.e., it has the same dimensionality as the spectral radiance vector.

The logit transform allows us to work with a quantity, \( z \), which is defined in the range \([-\infty, +\infty]\) and, hence, transform the emissivity, which is defined over the range \([0,1]\), to a new unbounded parameter. Once back transformed, the retrieved emissivity is forced to be a number in the interval \([0,1]\). Because of the logit transform also avoids the retrieval of un-physical emissivities and/or the use of boundary constraints on each parameter.

Using the chain rule for the derivative of a composite function, we can write the derivative with respect to \( z \) in terms of the derivative with respect to emissivity,

\[
d_2(i) = \frac{\partial R(i)}{\partial z(i)} = \frac{\partial R(i)}{\partial \epsilon(i)} \left( \frac{\partial z(i)}{\partial \epsilon(i)} \right)^{-1} = \frac{\partial R(i)}{\partial \epsilon(i)} \epsilon(i)(1 - \epsilon(i)) \quad i = 1, \ldots, M \quad (9)
\]

Note that at this stage the emissivity derivative is a scalar function which depends on the wave number alone. Based on \( d_2 \), we can easily linearize the forward model with respect to \( z \). The component pertinent to emissivity reads,

\[
\frac{\partial R(i)}{\partial z(i)} (z(i) - z_o(i)) = d_2(i)(z(i) - z_o(i)) \quad i = 1, \ldots, M \quad (10)
\]

with \( z_o \) a suitable first-guess point. In vector notation Equation (10) can be written according to
\[ \mathbf{K}_{z,2}(z - z_0) \] 

with \( \mathbf{K}_{z,2} \) a diagonal matrix with elements, \( K_{z,2}(i,i) = d_2(i) \) and \( z, z_0 \) the vectors with elements \( z(i), z_0(i) \), respectively. Second, we develop the \( z \) spectrum in a truncated PCA series, with truncation point, \( \tau \leq M \). To this end we need a suitable ensemble of emissivity spectra to develop the PCA basis. For the present analysis, the appropriate emissivity ensemble is obtained by the ASTER (Advanced Spaceborne Thermal Emission Reflection Radiometer) Spectral Library version 2.0 [56] and the MODIS (Moderate Resolution Imaging Spectrometer) UCSB (University of California, Santa Barbara) Emissivity Library (http://www.icess.ucsb.edu/modis/EMIS/html/em.html, accessed on 19 June 2018). The ensemble counts 134 emissivity spectra (see Figure 6), which are representative, at a global scale, of senescent and green vegetation, bare soil, desert sand and rock emissivities. The emissivity spectra are laboratory measurements at a spectral sampling of 2 cm\(^{-1}\). They have been linearly re-sampled at the IASI sampling of 0.25 cm\(^{-1}\) and re-interpolated to the IASI range 645 to 2760 cm\(^{-1}\). The mean and standard deviation of the ensemble are shown in Figure 7.

**Figure 6.** Ensemble of laboratory emissivity spectra used to construct the PCA basis.

**Figure 7.** Mean and standard deviation of the ensemble of emissivity shown in Figure 6.
Before performing the PCA development, the quantity $z$ is first centered and standardized according to

$$\tilde{z}(i) = \frac{z(i) - \mu_z(i)}{s_z(i)} \quad i = 1, \ldots, M$$

with $\mu_z(i)$ and $s_z(i)$ the mean and standard deviation of the ensemble of the logit transform of surface emissivity spectra used to build up the PCA basis (see Figure 7). Let $U$ be the PCA orthogonal basis (of size $M \times M$), then we have

$$\tilde{z} = U_{\tau} c_{\tau}$$

with $c$ is the PCA scores vector, which can be computed according to

$$c = U^t z$$

where the superscript $t$ indicates the operation of matrix transposition. In the above equations $U_{\tau}$ is a matrix of size $M \times \tau$ formed with the first $\tau$ columns of $U$ and $c_{\tau}$ the vector $c$ truncated after the first $\tau$ elements.

Defining the diagonal matrix, $S_b$ according to

$$S_b(i,i) = s_z^2(i); \quad S_b(i,j) = 0; \quad \text{for} \quad i \neq j; \quad i,j = 1, \ldots, M$$

and considering the truncated expansion of $\tilde{z}$, the emissivity linear term of Equation (11) transform to

$$K_{z,2}(z - z_o) = K_{z,2} S_b^{\frac{1}{2}} U_{\tau} (c_{\tau} - c_{0,\tau})$$

or by simply setting $K_2 = K_{z,2} S_b^{\frac{1}{2}} U_{\tau}$

$$K_2(c_{\tau} - c_{0,\tau})$$

which shows that the emissivity state vector ($v_2$ in Equation (6)) contains the PCA coefficients or scores, $v_2 \equiv c_{\tau} = (c_1, \ldots, c_\tau)$. In this way, we have $N_2 = \tau$ so that $K_2$ is a matrix of size $M \times \tau$, $v_2$ is a vector of size $\tau$. Thus, if we use $\tau << M$ we can achieve a large dimensionality reduction and represent the whole emissivity spectrum over the IASI spectral coverage from 645 to 2760 cm$^{-1}$ with a few PCA scores.

To select the appropriate truncation point $\tau$, we look at the eigenvalues of the PCA decomposition. These eigenvalues are shown in Figure 8 and drops below 1 at $\tau = 20$. Since we work with standardized quantities, values below 1 mostly indicate noise, therefore we choose $\tau = 20$ as truncation point. This choice agree with the so called Kaiser criterion (e.g., [57]) and corresponds to an explained variance of 99.04%. It should be stressed that with twenty PC scores we can represent all the important spectral features. The infrared emissivity spectrum is a smooth function of wave number and mainly in case of desert sand or soil containing quartz, there is a sharp spectral feature around 8.6 $\mu$m due to the reststrahlen band. With twenty scores this feature is fully resolved as seen in Figure 9, which shows a typical example of desert sand emissivity and its reconstruction with 5, 10 and 20 PCA scores. In general, the emissivity of natural and terrestrial materials in the infrared can be fully resolved at a spectral resolution of 2 cm$^{-1}$ (e.g., [56]), therefore 20 PCA scores should be enough to correctly represent emissivity of Earth’s surface.
Figure 8. Singular values of the PCA decomposition of the ensemble of spectra shown in Figure 6.

Figure 9. Example of PCA reconstruction of a typical desert sand emissivity with a number of PC scores equal to 5, 10 and 20, respectively.

2.3.3. Implementation and Emissivity Error Analysis

According to Optimal Estimation [44], for the practical implementation of the retrieval approach we have to choose the background (both state vector, \( \mathbf{v}_a \) and associated covariance matrix, \( \mathbf{S}_a \)), the First Guess state vector, \( \mathbf{v}_o \), around which to linearize the forward model, and the observational covariance matrix, \( \mathbf{S}_\epsilon \), that is the matrix which models the noise affecting radiances. For \( \mathbf{S}_\epsilon \) we use the full level 1C covariance matrix for IASI (see e.g., [53]).

We stress that the model \( \delta \)-IASI performs the mathematical inversion of the whole IASI radiance spectrum (8461 channels) to simultaneously retrieve the state vector, which includes the surface temperature (\( T_s \)) and emissivity (\( \epsilon \)), the atmospheric profiles of temperature (\( T \)), water vapour (\( Q \)), ozone (\( O \)), and HDO, and the columnar amount of \( \text{CO}_2 \), \( \text{N}_2\text{O} \), \( \text{CO} \), \( \text{CH}_4 \), \( \text{SO}_2 \), \( \text{HNO}_3 \), \( \text{NH}_3 \), \( \text{OCS} \) and \( \text{CF}_4 \). In a normal run, the profile quantities are specified on a pressure grid of 60 layers spanning the whole atmosphere from 1100 to 0.005 hPa. Considering that we have \( \tau = 20 \) PC scores...
and 10 scalar parameters to be retrieved, we have that in a normal run the size of the state vector is \( N = 330 \), of which \( N_1 = 310 \) for the atmosphere (included the surface temperature) and \( N_2 = 20 \) for emissivity.

For a fast convergence of the retrieval scheme, the background or \textit{a priori} and first guess state vectors should be chosen so that they are as much as possible in the linear region of the radiative transfer equation. Towards this objective, we use ECMWF (European Centre for Medium range Weather Forecasts) analysis for \((T, Q, O)\) collocated in space and time with the IASI soundings (see [53] for details). The use of ECMWF \((T, Q, O)\) analysis is providing a strong constraint in the inverse scheme and we think that the corresponding profiles are reliable. This allows us to relax the constraints for the trace gases. In effect, their profiles are normally adjusted starting from a climatology as a first guess and the retrieved profiles reach their final values within one or two iterations.

For emissivity we use the mean value of the ensemble of states shown in Figure 6 (note that in the PC-space the mean corresponds to the zero vector) as background state. Based on the PCA theory, the covariance matrix of the PC scores is diagonal with elements equal to the singular values (see Figure 8). Thus, for the background matrix of the \( \tau \) PC scores we use a diagonal matrix of size \( \tau \times \tau \) with the diagonal elements set to the first \( \tau \) eigenvalues. Again, we stress that we use \( \tau = 20 \) PC scores.

Usually, the First Guess vector is set to the background state. However, this is not feasible for emissivity retrieval because the mean values (Figure 6) could be too far from the true emissivity state, and, therefore, could be outside the linear region around the true state vector. The consequence could be either that convergence is not reached or only after many iterations. To ensure we stay as close as possible to the linear region of the inverse problem solution, we use the University of Wisconsin (UW) Baseline Fit (BF) Emissivity database (UW/BFEMIS database, e.g., http://cimss.ssec.wisc.edu/iremis/ [12]) to set a suitable First Guess. The procedure to calculate the emissivity First Guess with spectral coverage and sampling matching that of IASI is described in [50]. It is important to stress UW/BFEMIS database is used to obtain a suitable emissivity First Guess and not to define the background state vector, which is still the mean emissivity spectrum shown in Figure 6.

A detailed account on background state vectors and related covariances matrices would take too long to explain here, at expense of the main topic of this paper, which is validation of the emissivity product. Thus, for the sake of brevity, the interested reader is referred to [53,54]. Once we have defined and/or computed background states and covariance, First Guess state and spectral radiances, the estimate \( \hat{v} \) of \( v \) according to Optimal Estimation is obtained by

\[
\hat{v} = v_a + (S_a^{-1} + K^T S_e^{-1} K)^{-1} K^T S_e^{-1} ((R - F(v_o)) - (K(v_a - v_o))
\] (18)

We stress again that \( \hat{v}, v_a, v_o \) are the parameter state vector (estimated), the \textit{a priori} or background vector and the first guess state vector, respectively; \( R \) is the radiance vector and \( F \) is the forward model; \( S_a \) and \( S_e \) are the background and observational covariance matrices, respectively.

For a full assessment of the accuracy and/or error affecting the estimate \( \hat{v} \), Equation (18) has to be complemented with the Averaging Kernels (e.g., [43]),

\[
\frac{\partial \hat{v}}{\partial v} = (S_a^{-1} + K^T S_e^{-1} K)^{-1} K^T S_e^{-1} K
\] (19) and the \textit{a posteriori} or retrieval covariance matrix,

\[
\hat{S}_c = \left( S_a^{-1} + K^T S_e^{-1} K \right)^{-1}
\] (20)

It has to be stressed that the \textit{a posteriori} covariance matrix \( \hat{S}_c \) applies to the atmospheric parameters and the PCA coefficients of the logit-transformed emissivity spectrum. To obtain the covariances in physical emissivity space, the above covariance matrix has to be adequately transformed.
The transformation is not straightforward. It can be obtained in two steps, which are here outlined for the benefit of the reader.

First, we need to consider that the logit-transformed emissivity, $\mathbf{z}$ is related to the PC scores through the transform

$$
\mathbf{z} - \mathbf{z}_0 = S^\top_2 \mathbf{U}_\tau (c_\tau - c_{0\tau})
$$

Thus, the transform from the whole state vector (atmosphere+emissivity) with $c_\tau$-elements to that with $\mathbf{z}$-elements reads

$$
\begin{pmatrix}
\mathbf{v}_1 - \mathbf{v}_0 \\
\mathbf{z} - \mathbf{z}_0
\end{pmatrix} =
\begin{pmatrix}
\mathbf{I}_1 & 0 \\
0 & \mathbf{A}_F
\end{pmatrix}
\begin{pmatrix}
\mathbf{v}_1 - \mathbf{v}_0 \\
c_\tau - c_{0\tau}
\end{pmatrix}
$$

where $\mathbf{I}_1$ is the identity matrix (size $N_1 \times N_1$) and $\mathbf{A}_F = S^2_2 \mathbf{U}_\tau$.

If we define

$$
\mathbf{A}_T =
\begin{pmatrix}
\mathbf{I}_1 & 0 \\
0 & \mathbf{A}_F
\end{pmatrix}
$$

then the covariance matrix corresponding to the $\mathbf{z}$-elements is given by

$$
\hat{\mathbf{S}}_z = \mathbf{A}_T \hat{\mathbf{S}}_c \mathbf{A}_T^\top
$$

It should be noted that this matrix has size $(M + N_1) \times (M + N_1)$, because the emissivity has been back projected to its physical space in which its dimensionality is that of the radiance vector which contains a separate emissivity value for each spectral channel.

Second, from the logit transform we learn that a variation in the emissivity at channel $i$, that is $\Delta \varepsilon(i)$ corresponds to a variation $\Delta \mathbf{z}(i)$ according to

$$
\Delta \varepsilon(i) = (1 - \varepsilon(i)) \varepsilon(i) \Delta \mathbf{z}(i)
$$

Let us define the diagonal matrix $\Delta$ according to

$$
\Delta_{ii} = \begin{cases}  
1 & i = 1, \ldots, N_1 \\
(1 - \varepsilon(i - N_1)) \varepsilon(i - N_1) & i = N_1 + 1, \ldots, M + N_1
\end{cases}
$$

then the final a posteriori covariance matrix, $\hat{\mathbf{S}}$ for the atmospheric and emissivity elements is given by

$$
\hat{\mathbf{S}} = \Delta \hat{\mathbf{S}}_c \Delta^\top
$$

Figure 10 shows the Averaging Kernels or AK (e.g., [43]) for the $\tau = 20$ PC scores used to represent the emissivity spectrum. The example applies to one of the IASI soundings shown in Figure 2. One important aspect of this AK is that it is nearly one at each PC score. The degrees of freedom are in fact 19.71, very close to the value of 20, which corresponds to a retrieval for which the twenty PC scores were fully resolved by the data. This is one of the main results of this study. In effect, Figure 10 shows that the emissivity spectrum (at least in its PC truncated decomposition) is determined solely from the data, i.e., the retrieval has no important dependence on the background. Hence, the smoothing error is zero and the dominant source of noise is the error contained in the spectral radiances and to a lesser extent the effectiveness of the PC truncated transform.
Figure 10. Example of Averaging Kernels for the twenty PC scores in the retrieved state vector. The example corresponds to one of the IASI soundings in Figure 2.

Figure 11 exemplifies the retrieved emissivity for one of the IASI soundings shown in Figure 2, whereas Figure 12 shows the accuracy of the retrieval computed according to Equation (27). It is seen that the retrieved emissivity is independent of the background and has an accuracy ranging from 0.1 to 1% throughout the spectral coverage from 645 to 2760 cm\(^{-1}\). The best accuracy corresponds to the window region of 8–12 \(\mu\)m (\(\approx 800\) to 1200 cm\(^{-1}\)), where the retrieval error is of order of 0.1%. In this range, the IASI radiometric noise is very small (\(\approx 0.15\) K at a scene temperature of 280 K), whereas at the end of the spectral range (beyond 2500 cm\(^{-1}\)), the noise can reach values as large as 4–6 K.

Figure 11. Example of emissivity retrieval corresponding to one of the IASI soundings in Figure 2.
3. Results and Discussion

Figure 13 shows the emissivity retrieval for the set of IASI soundings over the desert sand, whereas Figure 14 shows the retrieved emissivities for the set of soundings over the gravel plain (Compare Figure 2). As expected, the two sets show considerable differences in variability, which is more markedly seen in the spectral range of the reststrahlen bands ($\approx$ 1000 to 1250 cm$^{-1}$). In effect the scene is more homogeneous over desert sand than the gravel plain and the absorption of the reststrahlen band is mostly determined from the abundance and grain texture of quartz. The sands and dune-fields of the Namib desert have little vegetation cover, so that their main spectral features is due to the rich content of quartz, which for almost 70% has a grain size below 250 $\mu$m (40 cm$^{-1}$)) [58].
This abundance of fine grain size particles yields an emissivity spectrum with a characteristic fingerprint within the reststrahlen band at 8.6 µm (1162.8 cm$^{-1}$). Thus, the Namib desert sand soundings can be used as a benchmark to check the retrieval quality of the emissivity spectrum, in the sense that the emissivity spectrum must show this characteristic fingerprint. For fine grained quartz sand the reststrahlen band at 8.6 µm (1162.8 cm$^{-1}$) is characterized by two local minima (at 8.27 µm, 1209.2 cm$^{-1}$ and 9.32 µm, 1073 cm$^{-1}$) of which the deepest occurs at 9.32 µm (1073 cm$^{-1}$).

![Figure 14](image.png)

**Figure 14.** Retrieval of emissivity from IASI soundings over the gravel plain.

This structure tends to reverse with coarse-grained particulate (range of grain size 250 to 1500 µm) where the deepest minima occurs at 8.27 µm (1209.2 cm$^{-1}$), see, e.g., [59]. By looking at the slope of these two minima we can gain information about the grain size of the quartz sand. For the IASI soundings at hand the slope should be positive, which in fact is the case: Figure 15 compares the average emissivity for the soundings over sand and gravel plain. Similar results have been found by [16,23].

Finally, Figure 16 compares the mean emissivity of the six IASI soundings with the mean emissivity of the in situ measurements. The comparison includes the IASI soundings, which are best collocated with in situ measurements. The figure also shows the ±3σ interval as computed from the variability (standard deviation) of in situ measurements.

The comparison in Figure 16 shows that IASI is capable to retrieve the two strongest silicate fundamental vibration bands (800 cm$^{-1}$ and 1150 cm$^{-1}$), which occur in the 8–14 µm atmospheric window. The shape of the retrieved emissivity spectrum agrees well with the in situ measurements and differences are generally below 0.025. However, a bias of ≈ 0.02 exists between in situ measurements and IASI observations. This difference can be explained by the angle dependence of emissivity [60]. This dependence becomes important at zenith angles larger than 40°, where emissivity decreases with the increase of zenith angle. According to [60] the decrease is most pronounced in sand, which is rich in quartz, reaching a maximum of about 0.10 at 70°. Furthermore, the decrease depends on wave number and tends to be larger at ≈ 8–8.7 µm (1150–1250 cm$^{-1}$). The IASI soundings considered for the comparison have a zenith angle of: 37.04°, 52.12°, 31.76°, 25.41°, 52.28°, respectively. Another effect that has to be considered is the IASI diameter of the Field of View (FOV) that for the set of angles used here is larger than 12 km. Although flat and homogeneous, the target area shows emissivity variability at the scale of meters as shown by the variability of in situ observations.
Figure 15. Comparison of retrieved emissivity for desert sand and the gravel plain, which have been obtained by averaging the spectra retrieved for the soundings shown in Figures 13 and 14. The black straight line connecting the reststrahlen band minima has a positive slope as expected for fine-grained quartz sand.

Figure 16. IASI vs. in situ emissivity for the best collocated IASI soundings. The ± 3 σ interval is the variability (standard deviation) of in situ measurements.

The angular effect cannot be addressed with the set of five soundings used for comparison with in situ measurements, since they correspond to relatively large Field of View. However, if we consider the full set of IASI soundings presented in this paper (Figure 3), we can evidence the angular effect. The set has been divided in two sub-sets with observations corresponding to FOV < 15° and those with FOV > 25°. Actually, some 50% of the soundings within this latter set have FOVs larger than 40°. The emissivity retrieval corresponding to the two sets is shown in Figure 17 and as expected that with lower FOVs has larger emissivity. This exercise has been shown just to exemplify that our methodology is sensitive to angular effects. A more accurate assessment should consider the same target area seen at different angles, an exercise which is out the scope of the present analysis.
Emissivity has a dimensionality as large as that of the observed radiance vector, therefore for a given spectral observation, this usually leads to an inverse problem with more unknowns than data points. In this paper, the dimensionality issue has been handled by representing the emissivity spectrum with a truncated PC transform with twenty PCA coefficients. The PCA orthogonal basis was obtained from a set of laboratory measurements and the methodology was applied to IASI and a set of soundings over Gobabeb validation station (Namib desert), which was used to check the accuracy and sensitivity of the methodology.

To date, the methodology we present is the first fully physical scheme, which depends only on the data and exploits Optimal Estimation to retrieve twenty PCA scores. Once the twenty PCA coefficients have been retrieved, the emissivity spectrum is calculated by taking the back transform of the truncated PCA. Based on our previous work [23], the use of PCA requires fewer coefficients than the Fourier transform. Furthermore, the presented retrieval scheme is very fast because of the simplified algebra. In effect, the computational time of the inverse problem is almost negligible compared to the forward model. For a single IASI sounding the inverse procedure takes ≈ 1 s (with a processor Intel(R) Core (TM) i7-4712HQ CPU 2.30 GHz (Santa Clara, CA, USA), 16 GB Ram) to retrieve the complete list of parameters: surface temperature ($T_s$) and emissivity spectrum ($\varepsilon$), the atmospheric profiles of temperature ($T$), water vapour ($Q$), ozone ($O$), HDO, and CO$_2$, and the column amount of N$_2$O, CO, CH$_4$, SO$_2$, HNO$_3$, NH$_3$, OCS and CF$_4$. Another important aspect of our methodology is that atmospheric parameters are simultaneously retrieved with the surface parameters, therefore, the retrieval scheme does not require an atmospheric correction.

A set of in situ emissivity measurements over the spectral range from 750 to 1250 cm$^{-1}$ has been compared IASI emissivity retrievals: the results show very good agreement, particularly if the dependence on zenith angle and the inherent variability of emissivity are properly considered.

The fact that the methodology depends mostly on the data and much less on background constraints opens the way to effective change detection analysis with hyper-spectral sensors such as IASI and allows addressing important applications, e.g., the monitoring of ice sheet and desert extent, oil spills, forest fires and agriculture, to name but a few.
Author Contributions: G.M. performed radiative transfer calculations. The retrieval analysis was performed by C.S., S.V. developed the retrieval analysis. G.L. modified radiative transfer model to properly take into account the solar reflection. F.M.-G. and L.P. conceived the field experiment and executed the field measurements. The paper was written by G.M. and C.S.

Acknowledgments: IASI has been developed and built under the responsibility of CNES. It is flown onboard the Metop satellites as part of the EUMETSAT Polar System. The IASI L1 data are received through the EUMETCast near real time data distribution service. The in-situ emissivity spectra were obtained during an international LST field inter-comparison experiment (FICE) organised within the framework of the FRM4STS project funded by the European Space Agency (ESA). Figure 3 were reproduced from Google Earth. Work partly supported by MIUR Smart City project.

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

Abbreviations

The following abbreviations are used in this manuscript:

AVHRR Advanced Very High Resolution Radiometer
ECMWF European Centre for Medium Range Weather Forecasts
EUMETSAT European Centre for the Exploitation of Meteorological Satellites
ESA European Space Agency
FOV Field of View
IASI Infrared Atmospheric Sounder Interferometer
LSA Land Surface Analysis
MIUR Italian Ministry of Education, University and Research
NPL National Physical Laboratory
PCA Principal Component Analysis
SAF Satellite Application Facility

References


34. Theocharous, E.; Barker Snook, I.; Fox, N.P. Results from the 4th CEOS TIR FRM Field Radiometer Laboratory Inter-Comparison Exercise Part 1 of 4: Blackbody Laboratory Comparison, ESA Technical Report OFE-D100 (Part1); ESA: Noordwijk, The Netherlands, 2017.
35. Theocharous, E.; Barker Snook, I.; Fox, N.P. Results from the 4th CEOS TIR FRM Field Radiometer Laboratory Inter-Comparison Exercise Part 2 of 4: Laboratory Comparison of Radiation Thermometers, ESA Technical Report OFE-D100 (Part2); ESA: Noordwijk, The Netherlands, 2017.
46. Amato, U.; Antoniadis, A.; De Feis, I.; Masiello, G.; Matricardi, M.; Serio, C. Technical Note: Functional sliced inverse regression to infer temperature, water vapour and ozone from IASI data. *Atmos. Chem. Phys.* 2009, 9, 5321–5330. [CrossRef]


