Robust Fitting of Diurnal Brightness Temperature Cycle

[1]*G. Udahemuka,* [1]*F. van den Bergh,* [2]*B.J. van Wyk,* [2]*M.A. van Wyk*

[1]Remote Sensing Research Unit, Meraka Institute, CSIR, Pretoria 0001.
[2]French South African Technical Institute in Electronics, Tshwane University of Technology, Pretoria 0001.

gudahemuka@csir.co.za, fvdbergh@csir.co.za, vanwykb@tut.ac.za, vanwykma1@tut.ac.za

Abstract

Land surface temperatures (LSTs) can be approximated from brightness temperatures observed from satellites. Estimation errors between observed brightness temperatures and a brightness temperature model of a given pixel would provide information for a pixel concerned. Robust fitting of observed Diurnal Temperature Cycle (DTC) taken over a day of a given pixel without cloud cover and other abnormally conditions such as fire can give a data based brightness temperature model for a given pixel. In this paper, diurnal brightness temperatures received from the METEOSAT Second Generation (MSG) satellite were interpolated for missing data based on a model, and a performance test was performed by comparing a new approach based on robust modelling with previous algorithms implemented on MSG data: An algorithm based on pseudo-physical modelling of the DTC and an algorithm based on Reproducing Kernel Hilbert Space (RKHS) interpolator. The simulation results show that the new approach outperforms the previous used criteria, in the sense that the true nonlinear model is more often found.

Keywords: Model selection, Diurnal Temperature Cycle, METEOSAT Second Generation.

1. Introduction

Satellite images contain information that can be used in assessing different characteristic parameters of the earth. Meteorological satellites provide information on the meteorological factors of the earth and land surface temperature (LST) can be approximated from the brightness temperature supplied by these satellites. One of the geostationary meteorological satellites, METEOSAT Second Generation (MSG) has an instrument, Spinning Enhanced Visible and Infrared Imager (SEVIRI) which it uses to provide brightness temperature data with a temporal resolution of 15 minutes. The data can be used as one of the land surface parameters for early detection of land surface changes. The detection of the land surface changes could then be used for diverse applications such as the detection of fires [1]-[4].

The Diurnal Temperature Cycle (DTC) provides brightness temperature variations of a given pixel for a full day. Using MSG, it is possible to get 96 brightness temperature samples over 24 hours. Mainly, data received in the 3.9 μm band, which has a saturation level of 335 K, are used in fire detection, because of the band properties in approximately eliminating solar reflection from the earth or from the atmosphere to the satellite, and this implies a closer correlation between the LST and the brightness temperature measured in this band. The DTC can have anomalies due to total or partial cloud cover over a given pixel which causes the data contaminated with cloud unusable for LST determination. Other anomalies which can make a DTC model unstable can be the solar reflection, precipitation, different ground heights, land surface characteristics (e.g., the diurnal amplitude of the surface temperature is larger for dry bare soil than for dense transpiring vegetation as reported in [5]), land cover, and wind fluctuations. Weather fluctuations must also be considered during the derivation of a DTC model. The data contaminated with clouds can be viewed as missing data, and in this paper, a method is implemented to derive a DTC model and this serves in interpolating the missing data in DTC.

Previous work in modelling DTC, includes models based on the physical properties of DTC provided in [6] and [7] that was improved in [8]-[10]. In [6]-[8], the interpolation was automatically performed and Levenberg-Marquardt least-squares scheme was used for error minimization. The interpolation reduced the effect of cloudiness and ensured that the corrected LST was not based on outliers. In [8], a new parameter was included to the model presented in [6]. This parameter comes from the fact that the minimum temperature of a given day could be different from the minimum temperature of its preceding or following day. This pseudo-physical model, named the Cosine DTC model, was improved in [9] with the inclusion of a new parameter which is due to the observation of two different harmonics in DTC one for

the rising edge part of DTC and another one for the falling edge of DTC. The Nelder-Mead simplex optimization was used in the robust estimate of the parameters. In [10], less prior knowledge on parameters is needed for the cosine DTC model than in the work in [8] due to their findings that one of DTC parameter can be expressed in function of the remaining parameters. A model driven approach was implemented in [9] which uses Reproducing Kernel Hilbert Space (RKHS) interpolators to interpolate the missing data. The approach results in an improvement to the cosine DTC model in a mean square error sense. In modelling DTC, the optimum number of parameters must not exceed the number of observed data points [7] since the processing of the high volume of data obtained would be highly involved.

In this paper, a new method is proposed to model the DTC. The method has been used for robust matching [11] and it can be adapted to non-linear time series modelling. Brightness temperature time series are one the variety of non-linear time series.

The paper is organized as follows. Section 2 details the RKHS for the DTC modelling. In section 3, the robust fitting approach to DTC is described, simulation results are the focus of section 4 and section 5 presents a conclusion.

2. Reproducing kernel Hilbert space model

A Reproducing Kernel Hilbert Space (RKHS) [12] is a Hilbert space H, to which is associated an inner product $(\cdot,\cdot)$ and a kernel $K(\cdot,\cdot):\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ such that $K(t,\cdot)\in H$ for all $t\in\mathbb{R}$ and which has the reproducing property. This reproducing property is expressed as

$$\left(F(\cdot),K(t,\cdot)\right)=F(t)$$

for all $t\in\mathbb{R}$. The result of the reproducing property on H is that

$$\left(K(s,\cdot),K(t,\cdot)\right)=K(s,t).$$

Given the input-output training data set $\tau=\{t_i,f_i\}_{i=1}^{N}$ where $f_i=F(t_i)+\varepsilon_i$ are noisy measurements of some unknown function $F(\cdot):\mathbb{R}\to\mathbb{R}$, find the minimum norm approximation $\tilde{F}(\cdot)$ of $F(\cdot)$ in the RKHS H subject to the constraints $\left(F(\cdot),K(t_i,\cdot)\right)=f_i$. In [13] is shown that $\tilde{F}(\cdot)$ is of the form

$$\tilde{F}(\cdot)=\sum_{i=1}^{N_a}a_iK(\tilde{t}_i,\cdot), \qquad (1)$$

where $N_a\le N$ due to the presence of noise and the kernel centres $\tilde{t}_i$ are deduced from τ by means of data reduction scheme in [14] and a_i is an element of a column vector $\boldsymbol{a}$ given by

$$\boldsymbol{a}=\boldsymbol{G}^{\dagger}\boldsymbol{f}$$

where $\boldsymbol{f}$ is a column vector with elements f_i and the Gram matrix $\boldsymbol{G}$ is a matrix with elements $K(\tilde{t}_i,t_j)$. The superscript † denotes pseudo inverse.

For the DTC modelling Dirichlet kernel [14] was used and is given by

$$K(s,t)=\frac{\sin\left\{\left(n+\frac{1}{2}\right)u(s-t)\right\}}{\sin\left(\frac{1}{2}u(s-t)\right)}, \qquad (2)$$

where u is a dilation parameter and n is the harmonic number. The Dirichlet Kernel is widely used in optimal sampling of unknown function with limited number of harmonics [15], [16].

3. Robust fitting approach to DTC

3.1. Singular value decomposition

The algorithm presented here has been adapted from the work in [11] and uses singular value decomposition (SVD). Using the SVD of a matrix in computations, rather than the original matrix, has the advantage of being more robust to numerical error. The SVD also exposes the geometric structure of a matrix, an important aspect of many matrix calculations. Despite its usefulness, the use of the SVD is computationally expensive due to the linearly increase of number of computations with the number of dimensions. The SVD also operates on a fixed matrix; hence it is not amenable to problems that require adaptive algorithms.

SVD has many uses in the field of remote sensing, due to the inherent attributes associated with this technique, which include: key vector analysis, dimensional reduction, robustness and excellent noise reduction. Key vector analysis [17] is a technique which is primarily used for signal characterization and allows land-cover types to be classified. The increasing use and dissemination of hyperspectral data is causing many data analysis problems, with traditional classification techniques, due to the large number of channels. SVD can be used in such cases, to significantly reduce the dimensionality of data sets.

3.2. Algorithm

The algorithm presented here for robust fitting of DTC gives the possibility of modelling a DTC of a given pixel based on a set of training DTCs for that pixel. Given a set of training DTCs, a small set of basis DTCs can be constructed that characterises the variation in the training set and can be used to estimate any of the training DTC. Each DTC in a training set of p DTCs, forms a 1D vector and forms the column in $96\times p$ matrix $\boldsymbol{A}$. It is assumed that the number of training DTCs, p, is less than 96. SVD is then used to decompose the matrix $\boldsymbol{A}$ as

$$\boldsymbol{A}=\boldsymbol{U}\sum\boldsymbol{V}^T$$

where $\boldsymbol{U}$ is an orthogonal matrix representing the principal component directions in the training set, $\sum$ is a diagonal matrix with singular values $\sigma_1,\sigma_2,\ldots,\sigma_p$ sorted in decreasing order along the diagonal. The $p\times p$ orthogonal matrix $\boldsymbol{V}^T$ encodes the coefficients to be used in expanding each column of $\boldsymbol{A}$ in terms of the principal component directions. If the singular values σ_k, for $k\geq K$, are small, then, since the columns of $\boldsymbol{U}$ are orthonormal, a new vector $\boldsymbol{e}$ can be approximated as

$$\boldsymbol{e}^*=\sum_{i=1}^{K}c_iU_i$$

where the c_i are scalar values and these values are calculated by taking the dot product of $\boldsymbol{e}$ (any input DTC) and the column U_i. This is a projection of the input DTC, $\boldsymbol{e}$, onto the subspace defined by the first K basis vectors. This approximation corresponds to the least-squares estimate of the c_i. c_i's give a reconstructed DTC that minimizes the squared error $E(\boldsymbol{c})$ between $\boldsymbol{e}$ and $\boldsymbol{e}^*$ summed over the entire DTC:

$$\begin{aligned}E(\boldsymbol{c})&=\sum_{j=1}^{96}\left(e_j-e_j^*\right)^2\\&=\sum_{j=1}^{96}\left(e_j-\sum_{i=1}^{K}c_iU_{i,j}\right)^2\end{aligned}\qquad(3)$$

To robustly estimate the coefficients $\boldsymbol{c}$, the quadratic error norm in Equation (3) would be replaced with a robust error norm ρ which was reported to have successful results in optical flow application [11],

$$\rho(x,\sigma)=\frac{x^2}{\sigma^2+x^2}$$

$$\frac{\partial}{\partial x}\rho(x,\sigma)=\psi(x,\sigma)=\frac{2x\sigma^2}{\left(\sigma^2+x^2\right)^2}$$

and minimize

$$E(\boldsymbol{c})=\sum_{j=1}^{96}\rho\left(\left(\boldsymbol{e}_j-\left(\sum_{i=1}^{K}c_iU_{i,j}\right)\right),\sigma\right)\qquad(4)$$

The value σ is a scale parameter that affects the point on the error function at which the influence of outliers begins to decrease. From the derivative of the error function, ψ-function, the rejection of outliers begins where the second derivative of ρ is zero and this means that those residuals where

$$\left|\left(\boldsymbol{e}_j-\boldsymbol{e}_j^*\right)\right|>\sigma/\sqrt{3}\qquad(5)$$

The residuals obtained by Equation (5) have reduced influence on the model data and can be viewed as outliers. DTC outliers are on those points of DTC affected for example by cloud.

The value of σ can be estimated from the data. The computation of the coefficients $\boldsymbol{c}$ involves the minimization of the nonlinear function in Equation (4). The minimization was performed using gradient steepest descent scheme with a continuation method that begins with a high value for σ and lowers it during the minimization. This operation results initially to no rejection of data as outliers and as the algorithm continues to iterate towards the local minimum, the influence of outliers is gradually reduced.

Given a robust fitting that recovers the principal structure in the input DTC, those points that were treated as outliers can be detected. If robust fitting results in an important number of outliers, then an additional fitting would be required by minimizing the error function

$$E(\boldsymbol{c})=\sum_{j=1}^{96}m_j\rho\left(\left(\boldsymbol{e}_j-\left(\sum_{i=1}^{K}c_iU_{i,j}\right)\right),\sigma\right)$$

where $m_j=\begin{cases}0 & \left|\left(e_j-e_j^*\right)\right|\leq\sigma/\sqrt{3}\\1 & \text{otherwise}\end{cases}$,

with $\boldsymbol{m}$ as an outlier vector, or "mask".

In this algorithm, $\boldsymbol{U}$ is calculated once during training and each image pixel on MSG full Earth disc can have its own orthogonal matrix $\boldsymbol{U}$ characterizing its brightness temperature variation over a day. Each time an input DTC

is observed, an inner product between the DTC and each column of $\boldsymbol{U}$ is calculated.

Using cosine DTC model and RKHS method, a pixel DTC model is obtained by using one observed DTC in the training stage. With this new method, the model can be trained by more than one observed DTC and this provides more information to the pixel DTC model.

4. Simulation results

The used data set is as reported in [9]: MSG data were for six different regions in South Africa and in each region, data for 5 different pixels were selected for the simulation. For each pixel, brightness temperature sequences of between four and six complete days (cycle) were collected from 2004-08-01 to 2004-08-29 and some segments spanning 4 hours of duration were removed to some cycles of a given pixel in order to simulate periods of cloud cover over the pixel. The dataset consisted of three different sequences for each pixel depending of where the removed segment was taken. The segment removal was performed either at time between sunrise and time of maximum temperature, at time after sunset or at 4 hours of maximum temperature of the observed DTC. Each removed segment consisted of 16 consecutive MSG samples. The removal of 16 consecutive samples which correspond to a period of 4 hours was chosen based on the results in [8] where it was reported that in case of cloud cover less than 4 hours over a pixel, the missed data due to the cloud cover can be well interpolated using the cosine DTC model reported in [8].

The Cosine DTC model reported in [8] and the RKHS approach were compared with this new approach. The Cosine DTC model in [8] and RKHS model were trained using a complete cycle with no missing data. Nelder-Mead simplex method was used to minimize the error between the cosine DTC model and the observed DTC.

In implementing RKHS, the kernel centres were placed at intervals of approximately 105 minutes (7 consecutive DTC samples), which gives an order of approximation N_a=14 in Equation (1). In Equation (2), the number of harmonics n=7 and the dilation u was the approximate of half-width of the kernel main lobe. After getting the model, the model was scaled and translated vertically to obtain the best least squares fit to any DTC of the pixel using Nelder-Mead simplex method. In implementing the robust fitting method, three DTCs without missing data were used in training.

The interpolation of missing data in DTC with missing samples was implemented using the three approaches, and the comparison of the approaches is based on mean square error (MSE) between the observed DTC and the DTC model.

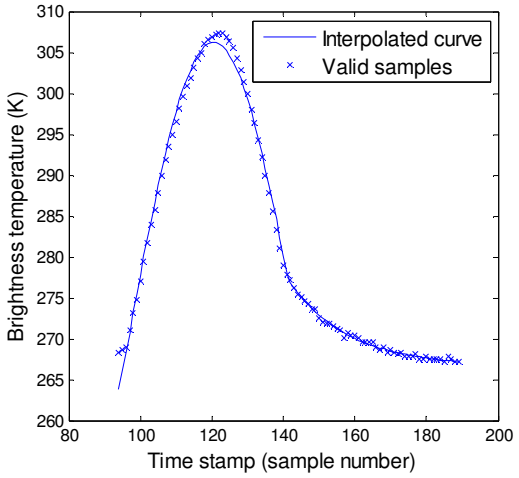

Figure 1: Sample MSG DTC with no missing data and fitted Cosine DTC model. MSE = 0.7465 and standard deviation of error = 2.2796.

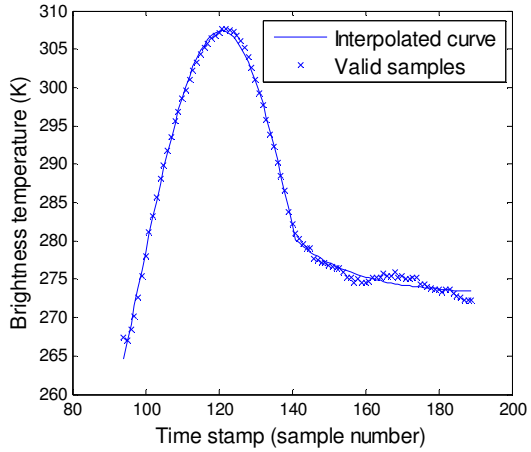

Figure 2: Sample MSG DTC with no missing data and fitted Cosine DTC model. MSE = 0.4727 and standard deviation of error = 0.9205.

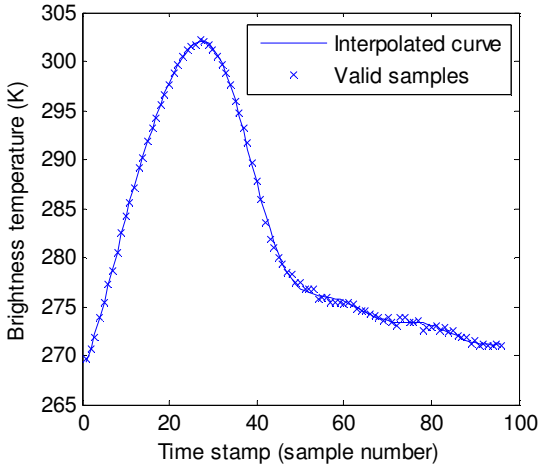

Figure 3: Sample MSG DTC with no missing data and fitted RKHS model. MSE = 0.0771 and standard deviation of error = 0.1392.

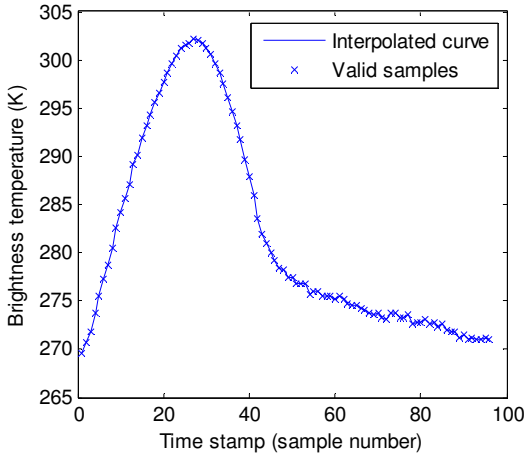

Figure 4: Sample MSG DTC with no missing data and fitted cycle using robust fitting. MSE is approximately zero.

In Figure 1, the Cosine DTC model does not fit at the point of maximum temperature of the observed DTC and in Figure 2, the Cosine DTC model does not follow humps in the observed data at the exponential decreasing part. In Figure 3 and Figure 4, MSG DTCs with no missing data and picked from the training set, are fitted to the model. As shown from the figures, both RKHS and robust fitting give better results at the point of maximum temperature and good results are also observed at the exponential decreasing part of the DTC where both approaches managed to adapt to the humps. The RKHS model exhibits a higher MSE compared with the model found using robust fitting. With robust fitting, the simulation exhibits MSE of close to zero.

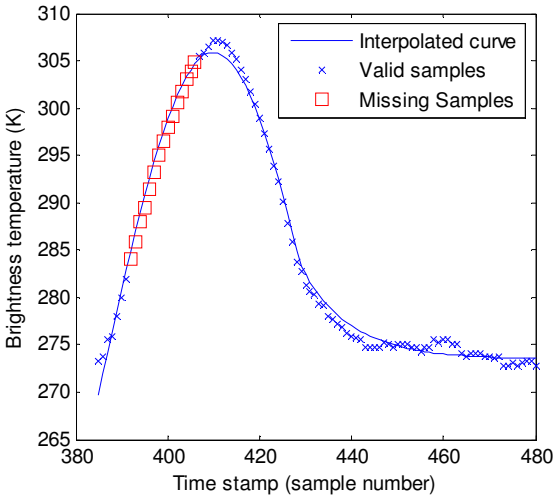

Figure 5: Sample MSG DTC with missing data and fitted cycle using cosine model. MSE = (0.8180, 0.1227).

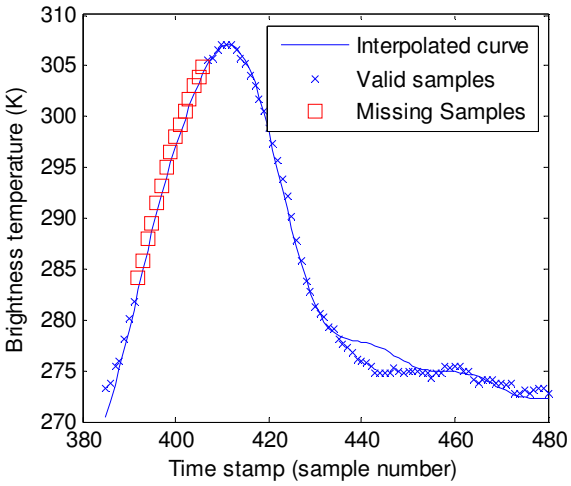

Figure 6: Sample MSG DTC with missing data and fitted RKHS model. MSE = (0. 5273, 0.1296).

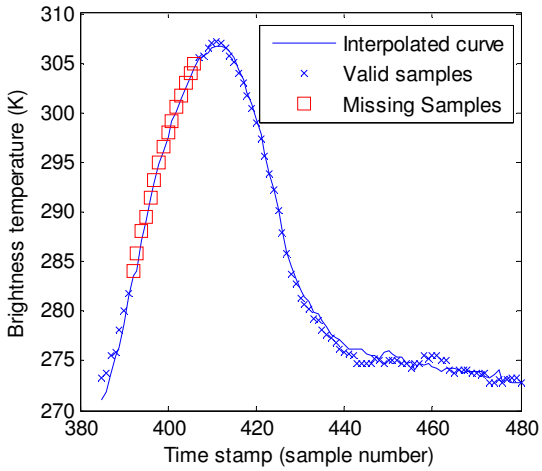

Figure 7: Sample MSG DTC with missing data and fitted cycle using robust fitting. MSE = (0.2243, 0.0502).

In Figure 5, Figure 6 and Figure 7, the interpolation of missing data was implemented to DTC with partial cloud cover and the robust fitting exhibits a low MSE compared to other methods and from the figures it can be seen that from the sunset, with the robust fitting, the model approximately approaches the observed data. The deviation of the model to the observations in Figure 7 is due to the fact that the DTCs used in training are different from the DTC observed. The interpolated curve exhibits the characteristic of the DTCs used in training and some jumps on the interpolated curve occur because of undetected anomalies in the observed DTC. Such anomalies are for example fire which could be the cause of the hump above the interpolated curve. As shown, RKHS model could not curve in the corner after sunset in Figure 6 and exhibits a higher MSE compared with the MSE found by interpolating using the robust fitting. The two MSE values in caption of Figure 5, Figure 6 and Figure 7 correspond to the measured MSE over all

Table 1: MSE obtained by fitting using cosine DTC modelling approach, RKHS approach and Robust fitting approach (standard deviation of error in parentheses).

Method	All samples	Missing samples only
Cosine DTC model	0.8878 (1.3968)	0.1334 (0.6292)
RKHS	0.5473 (0.6319)	0.1553 (0.3942)
Robust fitting	0.2635 (0.2769)	0.0663 (0.2823)

samples and the measured MSE over only the missing samples, respectively.

Table 1 shows an average in MSE and standard deviation of error over all cycles of a given pixel. From Table 1, a comparison was made on interpolation of a cycle with missing data. There is an improvement of above 50% in MSE in missing samples when compared robust fitting with other two approaches. Robust fitting also improve the time of execution by an average of 48% compared with the RKHS approach and compared with the cosine DTC model approach an approximately same time of execution was recorded

5. Conclusion

The presented results show the robustness of the method implemented in interpolating missing data in DTC with an improvement of approximately 50 % in MSE compared to RKHS interpolator and the Cosine DTC model. Future work will concentrate on a design of real time implementation of the modelled DTC with reduced parameters and in this case kernel methods for recursive least squares can be designed for the application. The brightness temperature would also be analyzed together with other quantities such as Normalized Difference Vegetation Index (NVDI), elevation, soil moisture and surface emissivity for the study of land changes to improve the DTC modelling.

6. Acknowledgement

The authors would like to thank EUMETSAT for making accessible MSG archived data which were used in this work. All MSG Data © 2007 EUMETSAT.

7. References

[1] A. Calle, J.L. Casanova, C. Moclan, A.J. Cisbani, E. Constantini, M. Zavagli and M. Greco, "Latest algorithms and scientific developments for forest fire detection and monitoring using MSG/SEVIRI and MODIS sensors," *Proceeding of 2nd International Conference on Recent Advances in Space Technologies*, Jun. 2005, pp. 118-123.

[2] F. van den Bergh and P.E. Frost, "A Multi-temporal approach to fire detection using MSG data," *Proceedings of Multitemp05*, Biloxi, Mississippi, May 2005, pp. 156-160.

[3] A. Calle, J.L. Casanova and A. Romo, "Fire detection and monitoring using MSG Spinning Enhanced Visible and Infrared Imager (SEVIRI) data," *Journal of Geophysical Research*, Vol. 111, 2006, pp. 1-13.

[4] G. Laneve, M.M. Castronuovo and E.G. Cadau, " Continous monitoring of forest fires in the mediterranean area using MSG," *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 44, No. 10, Oct. 2006, pp. 2761-2768.

[5] R. Nemani, L. Pierce and S. Running, "Developing satellite-derived estimates of surface moisture status," *Journal of Applied Meteorology*, Vol. 32, 1993, pp. 548–557.

[6] S. Schädlich and H. Fischer, "Determination of the diurnal wave of the land surface temperature with METEOSAT," *Proceeding of 9th Conf. on Satellite Meteorology and Oceanography*, American Meteorological Society, Boston, 1998, pp. 424–425.

[7] S. Schädlich, F.-M. Göttsche and F.-S. Olesen, " Influence of Land Surface Parameters and Atmosphere on METEOSAT Brightness Temperatures and Generation of Land Surface Temperature Maps by Temporally and Spatially Interpolating Atmospheric Correction," *Remote Sensing of Environment*, Vol. 75, 2001, pp. 39–46.

[8] F.-M. Göttsche and F.-S. Olesen, "Modelling of diurnal cycles of brightness temperature extracted from METEOSAT data," *Remote Sensing of Environment*, Vol. 76, 2001, pp. 337–348.

[9] F. van den Bergh, M.A. van Wyk and B.J. van Wyk, "A comparison of data-driven and model-driven approaches to brightness temperature diurnal cycle interpolation," *Proceeding of 17th Symposium of the Pattern Recognition Association of South Africa*, Parys, South Africa, Dec. 2006, pp. 252-256.

[10] G.-M. Jiang, Z.-L. Li and F. Nerry, "Land surface emissivity retrieval from combined mid-infrared and thermal infrared data of MSG-SEVIRI," *Remote Sensing of Environment*, Vol. 105, 2006, pp. 326–340.

[11] M.J. Black and A.D. Jepson, "EigenTracking: Robust matching and tracking of articulated objects using a view-based representation," *International Journal of Computer Vision*, Vol. 26, No. 1, 1998, pp. 63-84.

[12] N. Aronszajn, "Theory of Reproducing Kernels," *Trans. Am. Math. Soc.*, Vol. 68, 1950, pp. 337-404.

[13] D.G. Luenberger, "*Optimization by Vector Space Methods*," John Wiley & Sons, NY, 1969.

[14] M.A. van Wyk, "Hilbert Space Methods for Non-linear Function Approximation and Filtering," *Tech. rep*, CSIR/LEDGER, 2006.

[15] M.A. van Wyk and T.S. Durrani, "A framework for multiscale and hybrid RKHS-based approximators, " *IEEE Trans. on Signal Processing*, Vol. 48, No. 12, Dec. 2006, pp. 3559-3568.

[16] T. Strohmer and J. Tanner, "Fast reconstruction methods for bandlimited functions from periodic nonuniform sampling, " *SIAM Journal Numerical Analysis*, Vol. 44, No. 3, 2006, pp. 1073-1094.

[17] G. Herries, T. Selige and S. Danaher, "Singular value decomposition in applied remote sensing," *IET Colloquium on Image Processing for Remote Sensing*, London, UK, 13 Feb. 1996.