ELSEVIER

Early fire detection using non-linear multitemporal prediction of thermal imagery

A. Koltunov *, S.L. Ustin

California Space Institute (CalSpace) Center of Excellence, Center for Spatial Technologies and Remote Sensing, University of California, The Barn, One Shields Avenue, Davis, CA 95616, USA

Abstract

This paper presents a sub-pixel thermal anomaly detection method based on predicting background pixel intensities using a non-linear function of a plurality of past images of the inspected scene. At present, the multitemporal approach to thermal anomaly detection is in its early development stage. In case of space-borne surveillance the multitemporal detection is complicated by both spatial and temporal variability of background surface properties, weather influences, viewing geometries, sensor noise, residual misregistration, and other factors. We use the problem of fire detection and the MODIS data to demonstrate that advanced multitemporal detection methods can potentially outperform the operationally used optimized contextual algorithms both under morning and evening conditions.
© 2007 Elsevier Inc. All rights reserved.

1. Introduction and background

Early detection of fires from space-borne measurements is important from operational and economical perspectives, due to the need to monitor vast territories, the high rate of fire spread, substantially higher costs of fighting large fires, as well as profound consequences of biomass burning for climate, global carbon budgets, ecosystem functions, and other environmental costs. There are many factors currently limiting application of satellite remote sensing for early fire detection, among which are infrequency of passes, inevitable lag in data dissemination, and low spatial resolution. This places development of advanced data analysis methodologies that are capable of extracting critically needed information as a key element in justifying the need for new satellite missions.

Today's operational systems for satellite-based fire detection share the same general structure. They consist of several tests that are merged by an algorithm, coupled with auxiliary techniques to filter out false positives. The tests themselves, in effect, compare the observed intensity at an inspection (detection) image pixel with its expected intensity under the 'non-fire' hypothesis. While the structure and the algorithms for merging the tests are important and are always being improved, there is a significant motivation to develop additional and/or alternative tests that have comparable or better performance than the existing ones. An especially attractive situation is when the new and existing tests utilize different or independent information, leading to the possibility of boosting the detection performance by combining them. From this perspective, a multitemporal approach may have great potential, because the operational algorithms utilize only information from a single image. In this paper, we show how to construct novel anomaly tests by using non-linear multitemporal prediction of radiometrically uncalibrated thermal imagery and demonstrate their potential value for detecting small-scale fires. Below we first briefly discuss the ideas behind single-date and multitemporal detection.

1.1. Single-date approaches to fire detection

The single-date fire detection methods use fixed-threshold or contextual tests, or both. The fixed-threshold thermal tests (e.g.

---

* Corresponding author. Tel.: +1 530 754 4470; fax: +1 530 754 6353.
  E-mail address: akoltunov@ucdavis.edu (A. Koltunov).

Arino & Melinotte, 1998; Li et al., 2000) flag a pixel as fire if the brightness temperature in one or more spectral bands exceeds a pre-specified threshold. Because the optimal value of the threshold is spatially and temporally variable and unknown *a priori*, it is selected conservatively to reduce the false positive rate. As a result, only large flaming fires are detected with high probability.

Contextual tests typically attempt to predict the background intensity at a pixel in question by averaging the intensities (or brightness temperatures) across the neighboring pixels. The tests then decide that the pixel is a hot anomaly, if the observed value exceeds the predicted value plus ν times standard or mean absolute deviation of the neighboring pixel intensities. Thus, the performance of contextual tests is inversely related to the natural background variability in the inspected scene. This circumstance makes it more difficult to detect small-size fires during daylight when background temperature contrasts are relatively high. The contextual tests are used in the methods by Flasse and Ceccato (1996), Justice et al. (2002), Giglio et al. (2003), and others. A more detailed account of the strong and weak points of several fixed-threshold and contextual algorithms can be found in Ichoku et al. (2003).

1.2. Multitemporal detection

The multitemporal approach to detection defines fires as a class of anomalous *changes* in the scene.

In general, the dynamics of the environment and the sensor response function precludes detection of anomalous changes in the scene using simple subtraction of pixel intensity values across time. The difference images are highly cluttered, because the measured pixel intensities of different objects or materials undergo temporal changes in different and non-proportional amounts. Therefore, along with detecting the anomalous changes, a detection system would also create a high rate of false alarms. Several empirical unsupervised methods to analyze difference images and extract change regions are discussed in Melgani et al. (2000).

Instead of comparing pixel intensities, it is reasonable to retrieve and compare time-invariant quantities. In thermal infrared imagery, this entails estimating temperature and emissivity from calibrated thermal data. This is a complex problem that has not yet been completely resolved (Dash et al., 2002). Even if it had been resolved, much of the thermal information separating targets from background lies exactly in the temperatures *which are not useful*, because they themselves depend on *dynamic* weather conditions. Hence, the absolute value of the change in object temperature cannot always be used directly for anomaly detection.

Other researchers (Jensen, 1983; Schaum & Stocker, 1997) have observed that a multispectral image can be predicted by a linear operator of a previous multispectral image, provided the images are accurately registered, i.e. the pixel intensities can be transformed *jointly*. They applied this observation to detect changes in a scene and reported considerable improvements over the simple image subtraction. These methods for joint prediction share underlying ideas with the image normalization approach proposed by Schott et al. (1988) and extended by Furby and Campbell (2001) and by Du et al. (2002). In this approach, a set of pre-selected pseudoinvariant targets, such as roof tops, roads, parking lots, and other manmade objects are used to find the coefficients for a linear transformation between two images of different dates. The restrictive assumptions behind these methods are listed below:

1) images are registered across time to sub-pixel accuracy;
2) the pixel can be analyzed only if the observation from the previous time image is available for that pixel;
3) external factors, such as weather, illumination, and survey conditions, as well as preprocessing operations are spatially invariant;
4) internal properties of background objects are temporally invariant;
5) the observation process model is linear;
6) temporally invariant objects can be located using prior information.

Unfortunately, datasets obtained from space-borne platforms typically have features that violate these assumptions. In particular,

- both temporally and spatially variable viewing geometry, due to wide-angle surveillance;
- misregistration and flight line edge effects;
- spatially variable influence of weather conditions (e.g. partial cloud cover, local precipitation etc.);
- each image may have many pixels with missing values;
- impossibility or difficulty of excluding abnormally changed objects from the (pseudoinvariant) training set for background, including but not limited to:
  · target objects,
  · small convective clouds,
  · thin clouds typically bordering large clouds,
  · systematic and/or abrupt changes in vegetation (e.g. phenological stage transitions or harvest),
  · spatially and temporally variable sensor failures of non-extreme amplitude (e.g. stripes).

These factors limit the range of applications of simple multitemporal techniques for thermal data. To the best of our knowledge, multitemporal approaches have not yet been used for fire detection.

Nevertheless, Koltunov et al. (2003) suggested and theoretically demonstrated that the multitemporal prediction approach has the potential to overcome many of these problems. This resulted in a method for classification and anomaly detection under an assumption of dynamic environment conditions, called the Dynamic Detection Model (DDM), which generalizes the multitemporal approaches discussed above. They showed that the choice of past images used for prediction and the number of past images are important for anomaly detection performance. In particular, when the observation process is assumed linear, while both the temporal changes of object internal properties and the spatial variability of external factors are insignificant, then a minimum of eight past images are necessary for thermal imagery prediction.

In this paper (Section 2) we present a method for detecting thermal anomalies based on a modified version of the DDM using non-linear multitemporal prediction. At the training stage the method searches for a locally optimal set of past basis images. At the detection stage, these basis images are combined in a non-linear way to predict the background pixel intensities in the inspection image. Pixels whose observed intensity is statistically significantly different from the predicted intensity are flagged as anomalies.

Section 3 describes the application of the proposed method for MODIS (MODerate Resolution Imaging Spectroradiometer) for a thermal image sequence acquired over Northern California. The potential of the method for early fire detection is demonstrated in Section 4 by comparing the power of the corresponding tests constructed by the presented method, a linear DDM, the state-of-the-art contextual method, and a simple multitemporal technique-example: bi-date normalization.

2. Method: non-linear Dynamic Detection Model

2.1. Multitemporal non-linear model of background dynamics

Consider a scene S consisting of M pixel locations. The value of intensity measured for a given band at a spatial location s and a time moment t is denoted by $w(s,t)$. We also let W_t denote the band image collected at time t. The images in the sequence are obtained via a spatio-temporal observation process whose analytical form is unknown.

The background object intensity observed at a time t and location s can be represented as a function of intensities that were observed at s previously, at P past time moments. We term these moments *basis times*, and denote them $t_{\mathrm{b}} \stackrel{\text{def}}{=} (t_1, \ldots, t_P)$. Mathematically, the model is written as:

$$w(s,t) = \mathcal{H}[w(s,t_1), \ldots, w(s,t_P); \gamma(t)] + \varepsilon_t, \quad (1)$$

where $\gamma(t)$ stands for the parameter vector. This vector depends on the observation time t and does not depend on the spatial location. Hence, operator $\mathcal{H}$ is a space-invariant operator, i.e. it transforms past images into a current image *jointly*, following the same rule for all pixels. The images $W_{t_1}, \ldots, W_{t_p}$ will be called *basis* images and are all together denoted by W_{b}.

The variable ε accounts for the random part of the spatio-temporal variation in background intensity. We adopt the assumption that at each time moment ε is normally distributed with zero mean and unknown standard deviation σ_t.

The objects (pixels) whose intensity values at the detection time are observed to be significantly different from the ones predicted by the right-hand side of Eq. (1) are the objective of the analysis and called *anomalies*.

Readers interested in the general mathematical basis for why multitemporal predictability of physical spatio-temporal observations exists are referred to Appendix B.

2.2. Unsupervised model estimation

The functional form of $\mathcal{H}$ and the parameter vector γ are unknown and therefore must be estimated. As a feasible practical way of estimation we use a subset of a complete family (Milman, 1999) of simpler functions to approximate the unknown $\mathcal{H}$. In this paper we limit ourselves to seeking a quadratic model, which is the simplest non-linear model:

$$\mathcal{H}[W_{\mathrm{b}}; \gamma(t)] = \gamma_{00}(t) + \sum_{k=1}^{P} \gamma_{k0}(t) W_{t_k}$$
$$+ \sum_{k,\ell=1}^{P} \gamma_{k\ell}(t) W_{t_k} W_{t_l}. \quad (2)$$

Note that some coefficients in this representation can equal zero.

Given an inspection image W_t, the parameter vector $\gamma(t)$ can be estimated in the least squares sense, using *indicators* of prediction. By prediction indicators we term pixel locations $s_1, \ldots, s_N$, (N exceeding $P(P+1)/2 + P + 1$, which is the number of unknowns in model (2)), whose correspondence across time has been established. For example, the images of the scene can be registered across time or georeferenced. The estimated $\gamma(t)$ is used at the detection time to predict the intensities for all pixels for which observations at the basis times are available.

Thus, the principal sub-problems in our approach include

— determining the prediction coefficients γ at the detection time,
— choosing the basis time vector t_{b}, and
— determining the indicators of prediction.

2.2.1. Computing prediction coefficients

Given a vector of basis times t_{b} and a set of prediction indicators, the prediction coefficients are determined by stepwise regression. At each step of this method, terms are added to or removed from the model to end up with a maximal number of statistically significant terms. Along the way we obtain the estimate of σ_t (or equivalently, the r.m.s.e. of the predictor), which is adjusted for the degrees of freedom. Furthermore, the statistical significance $\zeta_{k\ell}$ of the estimated coefficients $\gamma_{k\ell}$ are determined. The coefficients whose significance values are less than a threshold α are zeroed and so is the significance. After the model parameters have been estimated, the indicators that are extreme outliers to the model (i.e. pixels for which the absolute value of the prediction residual exceeds $5\sigma_t$) are determined. They are removed from the set of indicators, and the coefficients $\gamma_{k\ell}$ are re-estimated. This process continues iteratively until no extreme outliers to the current prediction model are found.

2.2.2. Determining basis times

The basis times are determined in advance, at the stage of training the detection system. The training sequence is selected from all available past images of the scene. This sequence is split into Selection subsequence (SEL) and a Test subsequence (TST). Phenomenological considerations (Koltunov et al., 2003) suggest that both subsequences should span a high diversity of weather conditions and other conditions of data acquisition, in order to construct a prediction model that is appropriate for various objects and future external conditions. To reduce the influence of overall image brightness on basis

time selection, the images from TST are normalized to have the same mean and standard deviation.

In this work, selection of the basis time vector is addressed by forming and examining vector-candidates τ from SEL to minimize the prediction error function $E(\tau)$ over TST:

$$E(\tau) \stackrel{\text{def}}{=} \max_{t \in \text{TST}} [\sigma_t(\tau)] \Rightarrow \min_{\tau \in \text{SEL}}, \qquad (3)$$

subject to the requirement that all terms in the prediction model are statistically significant. In other words, among all τ-candidates we chose the one for which over the Test detection images the error of prediction was minimal.

Examining all candidates is not feasible computationally. Therefore, it is reasonable to examine a limited number of candidate-vectors to find a locally optimal one as follows. The Selection subsequence images are preprocessed to exclude from consideration images with too many missing observations. An initial $\tau = \tau_0$ is selected to include images acquired under essentially different weather and survey conditions, specifically, at times of the day spaced by at least 30 min. Among the basis time moments spaced by less than 30 min, the moment with the least number of missing observations is chosen. The rest of the moments are moved to the list of basis times to be added, denoted by τ_{add}. Next, we form the list of basis times to be removed from τ and denote this list by τ_{rem}. Initially, τ_{rem} is set to τ_0. Finally, the search algorithm described in Appendix A is run.

In general, the vector of basis times can be different for different bands.

2.2.3. Selecting the prediction indicators

The known locations that have anomalous or missing observations at any basis time are excluded from consideration. The indicators are selected at random from the remaining pixels.

2.3. Hot anomaly detection

At the detection stage, the inspection image W_t is first registered toward the basis image sequence. Next, the prediction coefficients $\gamma(t)$ corresponding to the inspection image are determined (Section 2.2.1), along with the r.m.s.e. of prediction, σ_t. For each pixel, the coefficients are introduced into Eq. (2) and the predicted value of intensity, $\widehat{w(s,t)}$, is computed by evaluating the right-hand side of Eq. (1). Finally, to detect hot anomalies the following test is applied:

$$\frac{w(s,t) - \widehat{w(s,t)}}{\sigma_t} > z. \qquad (4)$$

The threshold z controls the sensitivity of the detector. It can be fine-tuned experimentally to optimize the overall performance of a complete detection system, which is supposed to include the above test.

3. Application for MODIS thermal image sequence

The experimental work described in this Section has two objectives. The first objective is to evaluate the image prediction quality (goodness of fit) in real conditions using satellite remote sensing data from the MODerate Resolution Imaging Spectroradiometer (MODIS) instrument. Secondly, we compare the hot anomaly tests constructed by the non-linear DDM to the equivalent contextual tests currently used for fire detection, and to other methods.

3.1. Dataset

The scene $\mathcal{S}$ used in this study is located in Northern California and occupies over 150,000 km^2. The scene is confined within a bounding rectangle with the corner coordinates: -124.6W, 42.04N (upper-left) and -117.2W, 39.6N (lower-right). Fig. 1 displays the natural color image of the scene. The conditions in the scene are typical of the complexity of montane semi-arid regions in high fire risk western U.S. landscapes, with a mosaic of forests, shrubs, grasslands, urban structures, and water bodies. Vegetated and therefore burnable land is ubiquitous over the land area. The

Fig. 1. Natural color image of the test scene located in Northern California. The image size is 250×600 pixels at a ground resolution of 1 km.
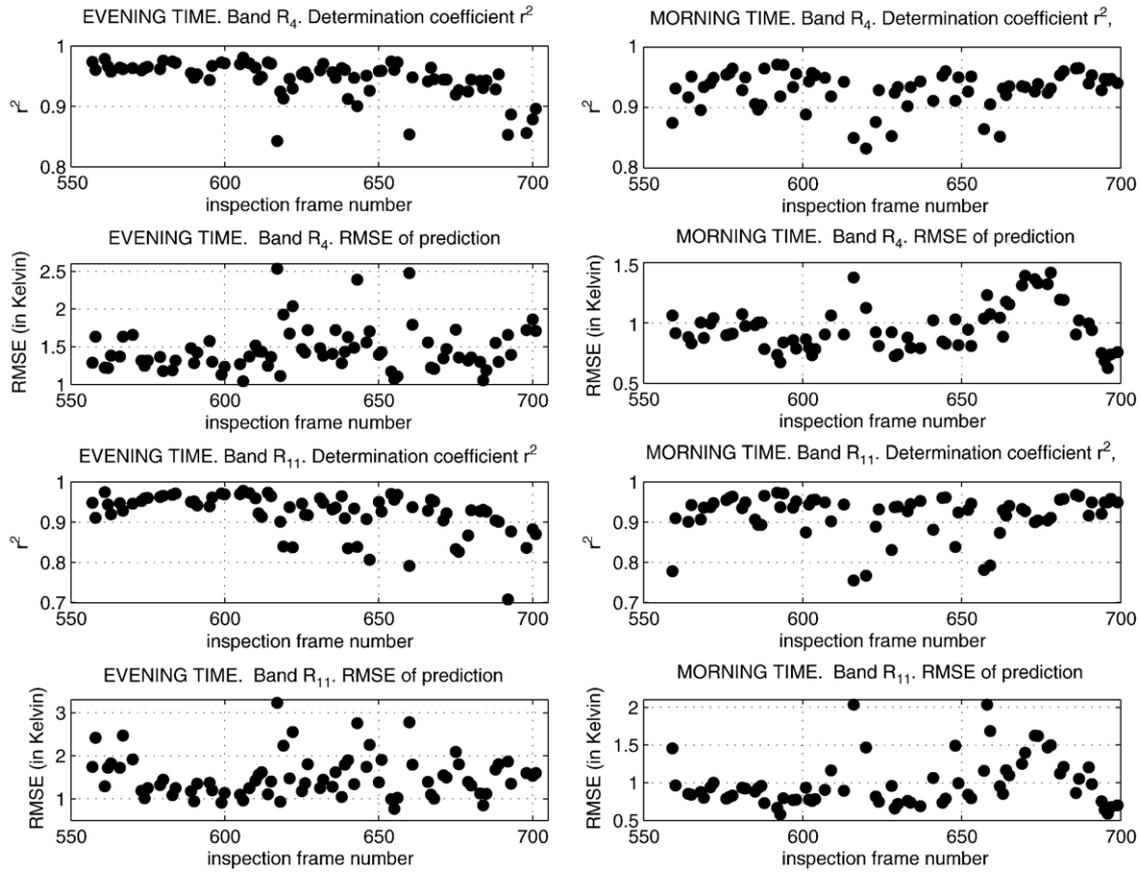
Fig. 2. Determination coefficient r^2 and the r.m.s.e. of background prediction, both adjusted for the degrees of freedom. The r.m.s.e. is calibrated to brightness temperature.

MODIS instruments on the Terra and Aqua satellites overpass $\mathcal{S}$ three to seven times each 24-hour period, collecting the data under different solar-view geometries and a nominal ground resolution of 1 km at nadir. Two thermal bands: band 22

($\sim$ 4 μm) and band 31 ($\sim$ 11 μm) available as part of the MODIS swath Level 1B standard products MOD021KM and MYD021KM, were used in this work. These products contain scaled integer values that could be linearly transformed to at-
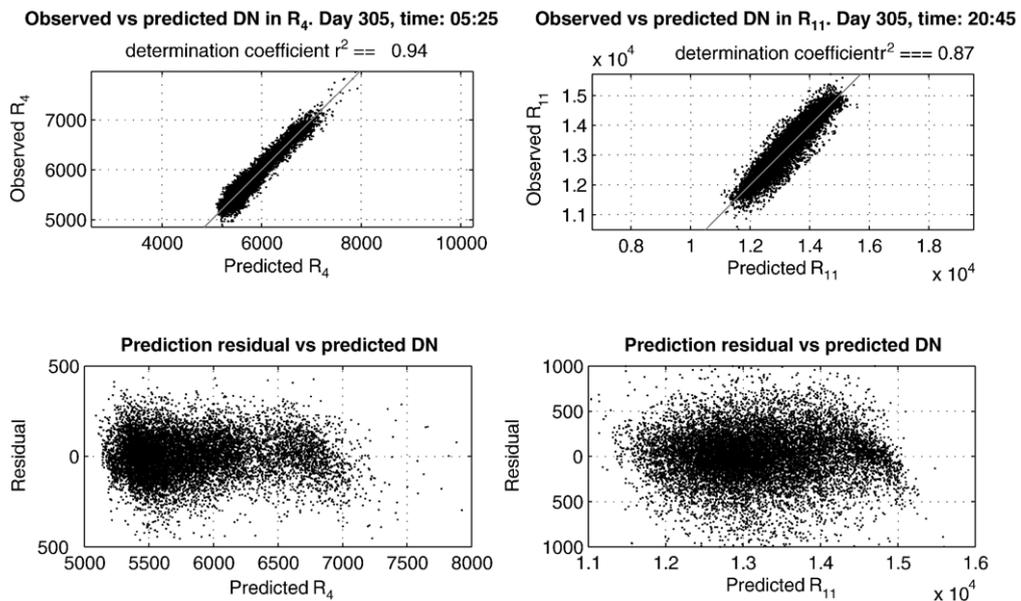
Fig. 3. Sample scatter plots of the image prediction residual.

sensor radiances. The scaled integer values for bands 22 and 31 will be denoted by R_4 and R_{11}, respectively, and the corresponding brightness temperatures — by T_4 and T_{11}.

The original multitemporal sequence consisted of 797 image frames acquired for eight months (March,1–November,1) of year 2004. The frames with the time-stamp differences not exceeding 5 min were mosaiced, because they are automatically part of the same data swath. The resulting sequence contained 701 images. The first 556 images were used for training the anomaly detection algorithm, and the last 145 images spanning 52 days — for assessing the algorithm performance.

3.2. Method implementation details

All images in the sequence were georeferenced to a common projected coordinate system with nearest-neighbor interpolation. The resulting ground pixel size was 1 km. To avoid processing cloud-contaminated observations, the MODIS cloud mask data (MOD35and MYD35 products) were used. The pixels flagged 'not-determined', 'confident cloud', or 'probably cloud' in these products, were marked as having missing values.

The unsupervised training procedure described in Section 2.2 resulted in a vector t_b containing 27 basis times. The basis times were selected to minimize the prediction error for band R_4 only. First 556 images formed the Selection subsequence and the next 19 images (557 to 571) — the Test subsequence.

As can be seen from Eq. (1), the multitemporal predictor is applicable to only those pixel locations that were available at all basis times. However, primarily due to spatially and temporally variable cloud cover, many pixels have missed observations in one or more basis images. Therefore, these pixels are not predictable using the best basis time vector resulting from the training procedure. To overcome this problem, a straightforward procedure was implemented in this paper. Given the best basis time vector t_b of length P, at the training stage we constructed P additional vectors of length $P-1$ each, by excluding one basis time from t_b. At the detection time, the prediction coefficients were computed for each vector as described in Section 2.2.1, using as indicators at most $2 \cdot 10^4$ pixel locations. These locations were chosen at random out of those that were not missing in the current basis images or in the inspection image. This procedure allowed us to predict ~ 70–80% of the pixels available at the detection time, which is about 10^5 pixels per
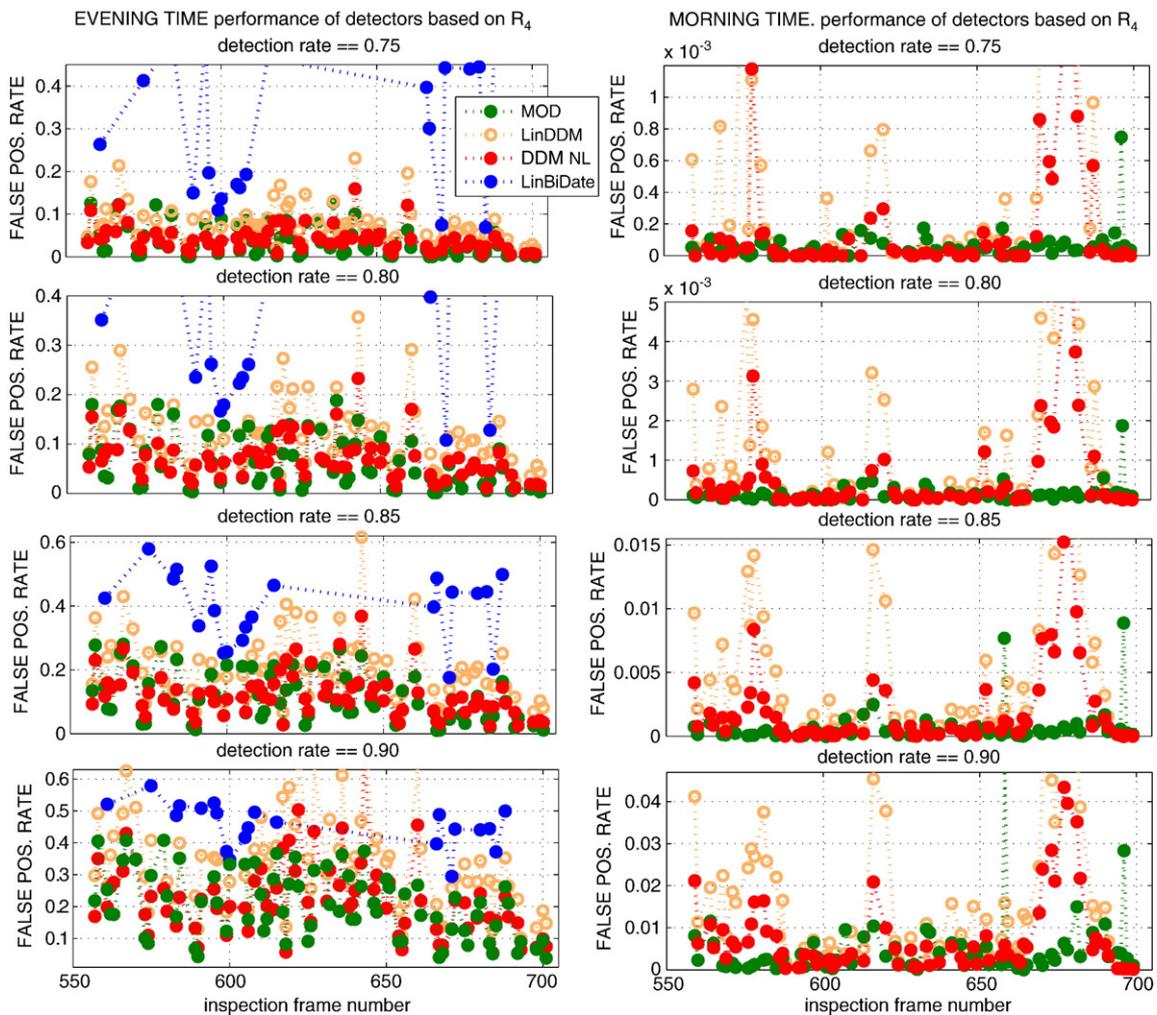
Fig. 4. False positive rates at several fixed true detection rates for detection algorithms applied to the 4-μm band R_4 only.

inspection image. Finally, for each pixel a predictor with the smallest r.m.s.e. (adjusted for the degrees of freedom) is chosen among the predictors for which the pixel does not have missing values at the basis times.

The statistical significance threshold α was set to 3.5.

4. Performance evaluation

4.1. Goodness of fit

Fig. 2 displays the sequences of determination coefficients r^2 and the sequences of the r.m.s.e. values calibrated to brightness temperature. Note that the values of r^2 are high for both thermal bands, R_4 and R_{11}, both evening and morning times. The background prediction performance does not deteriorate with the range of prediction, which manifests itself in the temporal consistency of r^2 and r.m.s.e. The determination coefficients are similar between the evening and morning images. Fig. 3 plots the prediction residuals for typical example-images. Note that the background pixels included many undetected cloud pixels, often flagged 'probably clear' in the MODIS cloud mask product. These pixels are in fact 'cold' outliers to the model. Therefore they considerably influence the quality of the background model.

4.2. Simulated fires on real background

This Section describes the comparative experimental evaluation of the potential of the presented anomaly detection method in an example detecting small-scale fires. We have simulated 10 groups of 100 idealized fires each. All 1000 fires were assumed to have the same kinetic temperature — 600 K and placed in the pixel centers. The area of each fire in the k-th group was set to $k \cdot 100$ m^2. The emitted radiance (under the blackbody assumption) was computed by Planck's law for each fire to alter the actual pixel background radiance, according to the anomaly area proportion and the band wavelengths. Finally, the radiances were converted back to the scaled integer values. For each of 145 inspection images, the simulated fires were placed at random in those pixels that

a) were located on land
b) did not have missing observations (see Section 3.2)

The rest of pixels, excluding known actual fires or pixels having missing values, were considered 'non-fire' pixels.

Simulated fires on real backgrounds were used to compare the proposed method to other detection methods (Section 4.2.1).
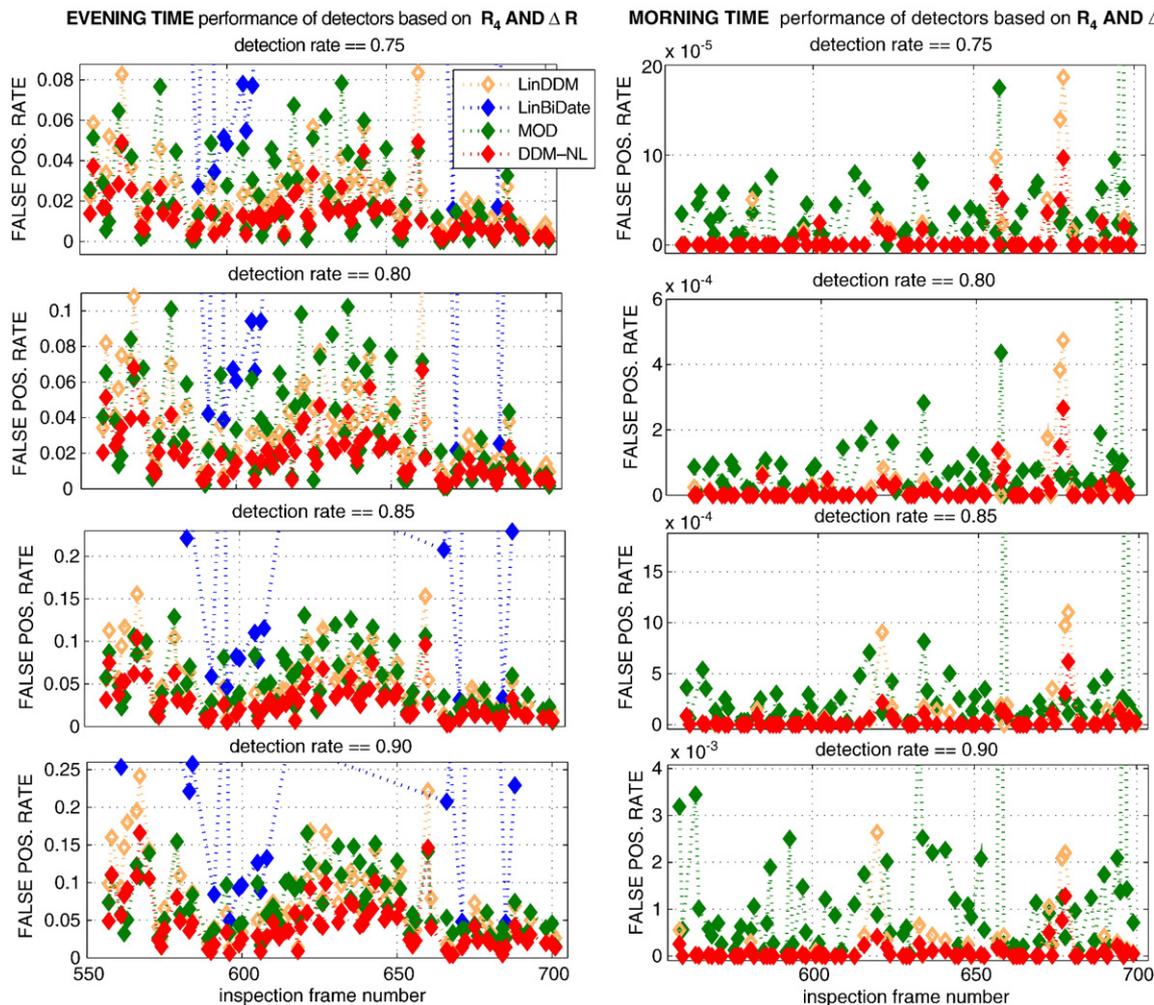
Fig. 5. False positive rates for the detection algorithms combining the R_4-test with the ΔR-test.

With these data we assessed the reliability of detecting a random new fire (ignition) that could occur anytime and anyplace on land and occupy an area of 100 to 1000 m^2 at the detection time. In general, using the simulated fire data may bias the *absolute* characteristics of the method performance. However, the use of simulated data is necessary and acceptable for *comparing* different methods or their components. Moreover, actual fires occur in the image with low probability. Therefore, modeling anomalies is often the only practical way to obtain a large and representative test set over the limited time range and scene size.

4.2.1. Comparison with alternative approaches

We assess performance of the proposed anomaly detection method (further abbreviated as DDM-NL) by comparing it with three other techniques, as abbreviated and described below.

4.2.1.1. MOD. The contextual method implemented in the standard MODIS fire product (MOD14) logically combines a series of contextual and absolute-threshold tests, preceded and followed by additional algorithms to filter out false positives. The contextual tests are applied to the swath data and use valid neighboring pixels in windows of variable size to estimate two background statistics: the mean and the mean absolute deviation. Details can be found in Giglio et al. (2003). In our comparison we used two contextual tests comprising the central part of the method:

$$T_4 > \mu_{T_4} + v_4 \delta_4 \tag{5}$$

$$\Delta T > \mu_{\Delta T} + v \delta_{\Delta T}, \tag{6}$$

where $\Delta T \stackrel{\text{def}}{=} T_4 - T_{11}$; $\mu_{T_4}, \mu_{\Delta T}$ denote the mean values; and δ_4, $\delta_{\Delta T}$ denote the mean absolute deviations. The thresholds v and v_4 that can be calibrated to detection confidence control the test sensitivity. The optimized selection of valid neighbors and the window size was implemented as described by Giglio et al. (2003).

4.2.1.2. LinBiDate. A two-date method that predicts the detection image by linearly transforming a previous image taken at a time as close as possible to the same time of the day as the detection image, e.g.: $\widehat{R_4(t)} = a R_4(t_{\text{prev}}) + b$, and detects hot anomalies analogously to Eq. (4).

4.2.1.3. LinDDM. Multitemporal prediction by Eq. (1) with a linear operator of prediction.

The last two methods, LinBiDate and LinDDM, were included to assess the importance of two essential features of the presented method: the plurality of past images used in modeling background and the non-linearity of the model.

All methods but LinBiDate were applied to 145 detection times. To avoid inventing a way to deal with missing observations that would be appropriate for LinBiDate, we applied this method to 21 cloud-free evening images. The multitemporal methods were applied to R_4 and $\Delta R \stackrel{\text{def}}{=} R_4 - R_{11}$, which provided the equivalent for the brightness temperature tests, Eqs. (5) and (6) by Giglio et al. (2003). Because the

images were resampled with nearest-neighbor interpolation *before* simulating the anomalies, only pixels corresponding to the original swath data pixels were included in the background characterization by MOD, as is described in Giglio et al. (2003). Finally, for all methods we logically combined the two tests, the one based on R_4 and the one based on ΔR, to obtain the higher performance fire detectors. Note that all methods combined the single-band tests in the identical way: "a pixel is flagged fire if and only if both tests return true". Therefore, the discrepancy between the detection results were solely due to differences between methods.

Figs. 4 and 5 plot the false positive rates given the fixed true detection rates for 4-μm band test and for the combination of the R_4-test and ΔR-test, respectively. The sensitivity thresholds of each algorithm were adjusted to achieve the desired detection rates at each frame. The 4-μm band tests constructed by our method, DDM-NL, and by the state-of-the-art contextual method, MOD, have similar power and outperform other approaches in evening hours. For about 10% of the morning 4-μm images the multitemporal tests yielded remarkably more false positives than MOD. This can be attributed to local changes (relative to the basis times) in vegetation conditions, or to local sunlight influences that the multitemporal space-invariant background model failed to account for. However, in both
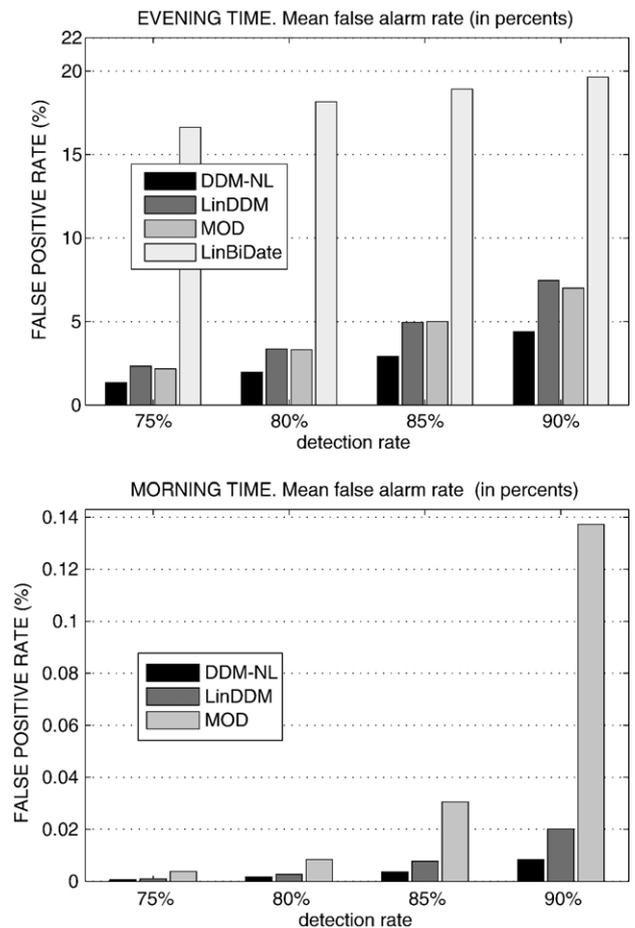
Fig. 6. False positive rates averaged over 52 days for the detection algorithms combining the R_4-test with the ΔR-test.

channels R_4 and R_{11} these anomalous changes were similar and therefore did not appear significant in ΔR. This phenomenon seems to be consistent. As a result, the detection algorithm DDM-NL using the combination of R_4-test and ΔR-test is noticeably better than its counterparts, as is evident from Fig. 5. Averaged over the 52-day period false alarm rates are shown in Fig. 6. Overall, DDM-NL committed about 60% fewer false positives than MOD in the evening and 500 to 1500% fewer false positives than MOD in morning hours.

We used fires of fixed absolute intensity, whereas the morning background was substantially cooler than evening. Hence, all methods performed for morning times better than for evening times.

5. Discussion

The objective of the prediction model is not to merely detect objects *with changed internal properties*. In fact, under real conditions after extended surveillance periods the object properties at most pixel locations are found to have changed since the basis times. When this is the case, the goal is to discriminate between two classes of changes: "natural" changes and anomalous ones. The natural changes are due to the combined effects of object internal properties and various external factors. For example, the soil thermal inertia depends on the soil water content, which in turn depends on the weather history, ground topography, soil chemistry, etc. In addition, natural changes can be attributed to the deviation that is not explained by the space-invariant prediction model, which has been derived from a physical perspective. In an unsupervised setting, pixel-examples sampled separately and representatively from the two classes are not available. However, we can and did make an assumption about the probability distribution of the magnitude of unexplained natural changes within the scene and estimate the parameter (σ_t) of that distribution.

In the current implementation, to overcome the problem of missing observations at basis times we excluded basis times one by one. An alternative method could be to *predict* the missing basis values by Eq. (1) with an alternative set of basis images. Both approaches (predicting basis observations and excluding basis times), actually build a plurality of predictors each of which is applied to a different, and in effect, random subset of the scene pixels. Although the currently implemented procedure is sufficient for the purpose of demonstrating the potential value of the method, it is necessary to derive optimal algorithms for missing observations at basis times. This will be an important subject of further research.

Formally, the method is scene-dependent. The issue of how much the detection results depend on the choice of the scene size needs further study. In practice, large monitored areas should be partitioned to separate scenes each of which could be processed in parallel.

The non-linear space-invariant prediction model is only a simplified approximation of the true spatio-temporal observation field. Therefore we inevitably introduced a prediction error. There are two valid and mutually related measures of the approximation accuracy: the r.m.s.e. and the spatial correlation

(clustering) of the prediction residual. The latter means that due to insufficiently sophisticated modeling the neighboring pixels tend to have similar values of prediction error. In this paper, we looked for a model that is locally optimal within a family of quadratic models. Future efforts should develop computationally feasible algorithms for selecting an optimal model out of broader families of model-candidates.

In fact, the first 19 images out of total 145 images used for testing were also involved in training the prediction model. Nevertheless, as apparent from Figs. 4 and 5 the detection performance is virtually the same, thus indicating that the model is not overfitted.

The fire test data used in assessing the method performance consisted of sub-pixel size anomalies only. For the presented method each pixel is processed separately. Therefore, the area of a single detectable fire is practically *unbounded from above*, contrasting with contextual methods. This feature may lead to improved detection of large-scale smoldering fires. Another advantage of the multitemporal approach is its applicability to sensors of various modalities. Therefore, future advances in multitemporal detection algorithms will improve detection systems of other sensors and platforms.

In the experiment described in this paper, the performance did not deteriorate with time. Therefore, it is reasonable to anticipate stable performance far beyond the 52-day prediction range we used. Nevertheless, anomalous changes accumulate over time. Therefore the same set of basis images will not remain good indefinitely long for all pixels. For example, as soon as an image region changed anomalously and permanently (e.g. burned areas), the past basis images can no longer be used for predicting the region. Thus, a future operational detection system based on multitemporal prediction will need to include an algorithm for updating the set of basis images.

In addition to issues already discussed, other technical issues must be addressed in the course of developing a future complete operational target detection algorithm, including:

— employing additional bands in the analysis;
— optimizing image registration of swath data;
— image pre-filtering to reliably mask out areas where the target object occurrence is unlikely;
— rejecting specific types of false alarms;
— developing an overall detection confidence measure;
— reducing computational costs of training and detection;
— statistical and phenomenological analyses of sub-algorithm performance by simulated and real data.

Also, following implementation of the first version of a complete algorithm, extensive tests using true real fire information and other sensor data must be performed.

Finally, we acknowledge certain limitations in the presented approach:

• the accuracy of the background prediction may be lower for areas with high cloud cover frequency. When the scene is severely clouded at the detection time, a sufficiently complex prediction model cannot be built for the lack of indicators. In

this case, the detection method may need to look for indicators beyond the scene being processed. Pixels that do not have even a single past cloud-free observation cannot be processed;

- non-linear multitemporal prediction is computationally more demanding than contextual algorithms, especially when the basis images are not cloud-free;
- severe misregistration of the images across time reduces the sensitivity of detection and may cause systematic false positive patterns.

6. Conclusion

The problem of fire detection from satellite observations has received considerable attention in remote sensing literature. The most prominent and optimized algorithms utilize spectral or spatial information, or both. Unfortunately, little progress, if any, has been reported about using temporal dimension of space-borne data for fire detection. In this paper, we presented a multitemporal algorithm for detecting hot anomalies. This algorithm, non-linear DDM, was applied to MODIS thermal image sequence and compared to other methods. As is mentioned in the previous section, we recognize that the current, non-optimized version of the presented non-linear DDM has not predicted the inspection images perfectly. Nevertheless, even at this early stage of development, the multitemporal prediction approach has demonstrated its productivity and usefulness for solving problems of real environmental complexity. Perhaps, it is more important to remember that the contextual detection methods and the Dynamic Detection Model are based on mutually independent information. Therefore, considerable gain in fire detection performance for the operational systems is anticipated after improving and intelligently merging the two approaches.

Appendix A. The search algorithm for basis times

Let $d(t,\tau)$ denote the distance from a time epoch t to a time vector $\tau=(t_1,...,t_P)$ and be defined by

$$d(t,\tau) = \min_{k=1,...,P} |t-t_k|. \tag{A.1}$$

Using the notations of Section 2.2.2 we describe the search procedure as follows:

Step 1. if both τ_{rem} and τ_{add} are empty, go to step 7.
Step 2. add to the current τ a new basis time $t_{k_{best}} \in \tau_{add}$, at which $\max_{k=1,...,P}[d(t_k,\tau)]$ is attained, and exclude it from τ_{add};

Step 3. if $E(\tau)$ decreases, accept current τ and go to step 1, otherwise: restore previous τ;
Step 4. if the list of basis time to be removed is empty, go to step 1;
Step 5. remove from the current τ the basis time $t_{k_{worst}} \in \tau_{rem}$ at which $\min_{k=1,...,P}[\max\{\zeta_k, \zeta_{k1}, ..., \zeta_{kP}\}]$ is attained, and exclude it from τ_{rem};
Step 6. if $E(\tau)$ decreases, accept current τ and go to step 4, otherwise: restore previous τ and go to step 1;
Step 7. quit searching.

Appendix B. Mathematical insight into the image predictability phenomenon

Below we mathematically derive the law of predictability of a spatio-temporal observation field. An example of an observation field is a remote sensing image of a scene. The prediction model will be derived using a set of simplifying assumptions that are necessary to emphasize the principle. As demonstrated in this paper by the MODIS thermal imagery prediction, the assumptions under which the model is applicable, are much weaker than those used in deriving the model. Theoretical explanation of this fact is given elsewhere (Koltunov et al., 2003).

Let $W=\{w_i\}_{i=1}^N$ denote a spatio-temporal field of physical observations, obtained via an observation process $\mathcal{F}$ (either known or not). Assume that for a given transform $\mathcal{F}$, an element (e.g. pixel intensity) w_i being observed at spatial location s at time t, depends on two types of factors: the internal properties (i.p.) of the observed objects, denoted by β, and the external influence factors (e.f.): $\mathcal{X}$. In real physical processes, both i.p. and e.f. depend on time t and spatial location s, and thus:

$$W(s,t) = \mathcal{F}[\beta(s,t); \mathcal{X}(s,t)]. \tag{B.1}$$

For the sake of brevity of the derivation, we introduce several constraints into the above general model (B.1).

(a) internal properties β do not change with time, i.e. $\beta(s,t)=\beta(s)$;
(b) external factors $\mathcal{X}(s,t)$ are the same for each element of W, i.e. $\mathcal{X}(s,t) = \mathcal{X}(t)$;
(c) each element w_i is autonomous, meaning that its i.p. are independent of the i.p. of any other w_j;
(d) any w_i has a finite number m_i of internal properties;
(e) the number of observation field elements is finite.

If constrained by (a)–(e), model (B.1) can be written in the separate-variable form as follows:

$$W(s,t) = \mathcal{F}[\beta(s); \mathcal{X}(t)]. \tag{B.2}$$

Measure the observation field at P time moments, *at which the external factors are essentially different*, thus obtaining a system of $N \cdot P$ equations

$$W_{t_j} \stackrel{def}{=} W(t_j) = \mathcal{F}[\beta; \mathcal{X}(t_j)],$$
$$\text{where} \quad j = 1,...,P; \quad P \geqslant \max_i \{m_i\}, \tag{B.3}$$

wherein each element w_i appears P times. Therefore, if operator $\mathcal{F}$ is invertible for the chosen set $t_1,\dots t_P$, then the i.p. can be uniquely found from Eq. (B.3) as

$$\beta = \mathcal{G}[W_{t_1}, \dots, W_{t_P}; \mathcal{X}_{t_1}, \dots, \mathcal{X}_{t_P}], \tag{B.4}$$

$$\text{where} \quad \mathcal{X}_{t_j} \stackrel{\text{def}}{=} \mathcal{X}(t_j).$$

Because Eq. (B.2) *always* holds, we can introduce Eq. (B.4) into Eq. (B.2) to obtain *at any new time*[1] t_v

$$W_{t_v} \stackrel{\text{def}}{=} W(t_v) = \mathcal{H}_{t_v}[W_{t_1}, \dots, W_{t_P}]. \tag{B.5}$$

The predictor $\mathcal{H}_{t_v}$ itself depends only on e.f. $\mathcal{X}(t_v), \mathcal{X}_{t_1}, \dots, \mathcal{X}_{t_P}$. Note that these factors should not be considered as only instantaneous influence factors. Some of the e.f. can be long-lasting, for example, some of the weather factors influencing temperatures. However, the predictor does not depend on internal properties that are different for different elements w_i of the observation field. Hence, this predictor transforms 'old' fields into a 'new' field *jointly*, element into element, according to the same law for all elements.

Thus, we have derived a property that can be conveniently formulated as follows:

I. Any physical field (including an optical or thermal image) currently observed can be arbitrarily accurately approximated by a space-invariant transformation of a finite number of fields previously observed.

The space-invariant form of the predictor remains intact even if the internal properties and external factors depend both on space and time, but these dependencies are separable.

References

Arino, O., & Melinotte, J. M. (1998). The 1993 Africa fire map. *International Journal of Remote Sensing*, *19*(11), 2019−2023.

Dash, P., Göttsche, F. -M., Olesen, F. -S., & Fischer, H. (2002). Land surface temperature and emissivity estimation from passive data: Theory and practice-current trends. *International Journal of Remote Sensing*, *23*(13), 2563−2594.

Du, Y., Teillet, P. M., & Cihlar, J. (2002). Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection. *Remote Sensing of Environment*, *82*(1), 123−134.

Flasse, O., & Ceccato, P. (1996). A contextual algorithm for AVHRR fire detection. *International Journal of Remote Sensing*, *17*(2), 419−424.

Furby, S. L., & Campbell, N. A. (2001). Calibrating images from different dates to 'like-value' digital counts. *Remote Sensing of Environment*, *77*(2), 186−196.

Giglio, L., Descloitres, J., Justice, C. O., & Kaufman, Y. J. (2003). An enhanced contextual fire detection algorithm for MODIS. *Remote Sensing of Environment*, *87*, 273−282.

Ichoku, C., Kaufman, Y. J., Giglio, L., Li, Z., Fraser, R. H., Jin, J. -Z., et al. (2003). Comparative analysis of daytime fire detection algorithms using AVHRR data for the 1995 fire season in Canada: Perspective for MODIS. *International Journal of Remote Sensing*, *24*(8), 1669−1690.

Jensen, J. R. (1983). Urban/suburban land use analysis. In R. N. Colwell (Ed.), *Manual of Remote Sensing*, Vol. 2 (pp. 1571–1666). Falls Church, USA: American Society of Photogrammetry.

Justice, C. O., Giglio, L., Korontzi, S., Owensa, J., Morisette, J. T., Roy, D., et al. (2002). The MODIS fire products. *Remote Sensing of Environment*, *83*(1), 244−262.

Koltunov, A., Koltunov, J., & Ben-Dor, E. (2003). Adaptive recognition under static and dynamic environment assumptions. In S. S. Shen & P. E. Lewis (Eds.), *Algorithms for Multispectral and Hyperspectral Imagery Proceedings of SPIE*, Vol. 5093 (pp. 601–612).

Li, Z., Nadon, S., & Cihlar, J. (2000). Satellite-based detection of Canadian boreal forest fires: Development and application of the algorithm. *International Journal of Remote Sensing*, *21*(16), 3057−3069.

Melgani, F., Moser, G., & Serpico, S. B. (2000). Unsupervised change detection methods for remote sensing images. *Optical Engineering*, *41*(12), 3288−3297.

Milman, A. S. (1999). *Mathematical Principles of Remote Sensing: Making Inferences from Noisy Data, Chapter 14* (p. 390). Chelsea, MI: Sleeping Bear Press.

Schaum, A., & Stocker, A. (1997). Subclutter target detection using sequences of thermal infrared multispectral imagery. In S. S. Shen & M. R. Descour (Eds.), *Algorithms for Multispectral and Hyperspectral Imagery Proc. SPIE*, Vol. 3071 (pp. 12–22).

Schott, J. R., Salvaggio, C., & Volchok, W. J. (1988). Radiometric scene normalization using pseudoinvariant features. *Remote Sensing of Environment*, *26*, 1−16.

---

[1] More precisely, for any new external conditions.