



→ AGRICULTURE

## Sentinel-2 Agriculture

---

### Benchmarking of Vegetation Status Algorithms

---



La force de l'innovation



## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Algorithm description</b>	<b>2</b>
2.1	LAI multi-linear regression from <i>in situ</i> data (MLRISD) . . . . .	2
2.2	LAI non linear inversion of a radiative transfer model (NLIRTM) . . . . .	2
2.2.1	Software issues . . . . .	3
2.2.2	Algorithm setup . . . . .	4
2.3	LAI on-line multi-temporal retrieval (OLMTR) . . . . .	4
2.4	Complete season LAI multi-temporal retrieval (CSMTR) . . . . .	5
2.5	Temporal metrics from the NDVI temporal profile . . . . .	6
2.5.1	Definitions . . . . .	6
2.5.2	Estimation . . . . .	7
2.5.3	Attempts to estimate the emergence date in real time . . . . .	7
<b>3</b>	<b>Quality criteria</b>	<b>8</b>
3.1	<i>In situ</i> data preparation . . . . .	8
3.1.1	<i>In situ</i> data reliability issues . . . . .	9
3.2	Evaluation criteria . . . . .	9
3.2.1	Initially proposed metrics . . . . .	9
3.2.2	Time profile score for the LAI estimation . . . . .	10
3.2.3	Criteria for the NDVI metrics . . . . .	10
<b>4</b>	<b>Benchmarking results</b>	<b>10</b>
4.1	MLRISD . . . . .	10
4.2	NLIRTM . . . . .	11
4.3	Multitemporal estimations . . . . .	11
4.4	Temporal metrics . . . . .	13
<b>5</b>	<b>Conclusions on algorithm selection</b>	<b>14</b>
5.1	LAI retrieval . . . . .	14
5.2	Temporal NDVI metrics . . . . .	15
<b>6</b>	<b>References</b>	<b>16</b>
<b>7</b>	<b>Appendix</b>	<b>17</b>
7.1	Temporal profiles . . . . .	17
7.2	LAI images . . . . .	37

## 1 Introduction

Two different products are proposed as vegetation status indicators. The first one is a leaf area index (LAI) estimation and the second one is a set of metrics describing the temporal evolution of the vegetation. The LAI estimation is performed at the pixel level for every pixel and for every acquisition date. The temporal metrics are computed also at the pixel level, but only at the end of the season using all the images acquired.

Four different approaches for LAI retrieval have been selected as well as a set of temporal metrics which provide information about the phenology of the crops.

The rationale for the choice of the 4 LAI retrieval approaches is the following:

1. One approach which can be applied when *in situ* field measurements are available, usually in small quantities.
2. One approach which can be used without *in situ* data, but using vegetation optical properties and radiative transfer models.
3. One approach which takes into account the temporal information and which can be applied at any point in time during the season and exploiting all the collected data up to that point.
4. One approach which exploits the complete vegetation cycle and which may be more robust than the previous one, but can only be applied at the end of the season.

This document presents the results of the validation and benchmarking of the different algorithms used for the products mentioned above.

## 2 Algorithm description

### 2.1 LAI multi-linear regression from *in situ* data (MLRISD)

The idea here is to use the field data in order to build a regression model using the satellite reflectances as predictors.

$$\widehat{LAI} = \sum a_i \rho_i + a_{NDVI} NDVI \quad (1)$$

The  $a_i$  are the coefficients of the regression which are estimated using the field data, and the  $\rho_i$  are the satellite surface reflectances for each spectral band  $i$ . It has been found that adding the NDVI as predictor together with the reflectances improves the quality of the estimation, since non linear information is added to the otherwise linear the model.

The model can be improved by using several images in order to better estimate the coefficients of the regression, but in this case, the solar and satellite angles should be added to the model:

$$\widehat{LAI} = \sum a_i \rho_i + a_{NDVI} NDVI + a_{ss} \cos(\phi_{ss}) + a_{sat} \cos(\theta_{sat}) + a_{sun} \cos(\theta_{sun}) \quad (2)$$

were  $\phi_{ss}$  is the azimuth angle between the satellite and the sun,  $\theta_{sat}$  is the satellite zenithal angle and  $\theta_{sun}$  is the Sun zenithal angle.

In our case, we can use the data from all available images and estimate a model for SPOT4 and a model for Landsat8.

### 2.2 LAI non linear inversion of a radiative transfer model (NLIRTM)

This is the well-known BV-NET approach developed by Weiss et al. [WBL<sup>+02</sup>]. PROSPECT-5 [FFA<sup>+08</sup>] and SAIL-4 [JVB<sup>+09</sup>] are used to simulate surface reflectances for a wide range of vegetation parameters and these simulations are used to train a non linear regression algorithm for the estimation of one of the variables (LAI in our case) using the surface reflectances as predictors.

The main advantage of this approach is that no *in situ* data is needed for the calibration of the regression model. However, the inversion of the simulation models is an ill-posed problem, and therefore, the quality of the solution may be low. Added to that, the accuracy of the bio-physical models used for the simulation will strongly influence the quality of the results.

Using the same set of predictor variables, a non linear regression model for the estimation of the error on the LAI estimation itself can be built.

### 2.2.1 Software issues

Although the INRA EMMAH team kindly provides their Matlab implementation at no cost, CESBIO has chosen to implement its own version of the algorithm in order to make it available as free software to the remote sensing community. CESBIO's implementation is based on the Orfeo Toolbox, and is available at <http://tully.ups-tlse.fr/jordi/otb-bv>.

CESBIO's implementation is not a port of the original Matlab code but an independent version which uses different versions of PROSPECT (version 3 in INRA's code and version 5 in CESBIO's) and SAIL. Also, CESBIO's code implements several machine learning algorithms for the model inversion (Support Vector regression, Multi-linear regression and Random Forest regression are available together with the classical Neural Network regression).

**2.2.1.1 Theoretical performances** Figure 1 and table 1 present the performances of the algorithm when used to invert reflectance data obtained by simulation through PROSPECT and SAIL.

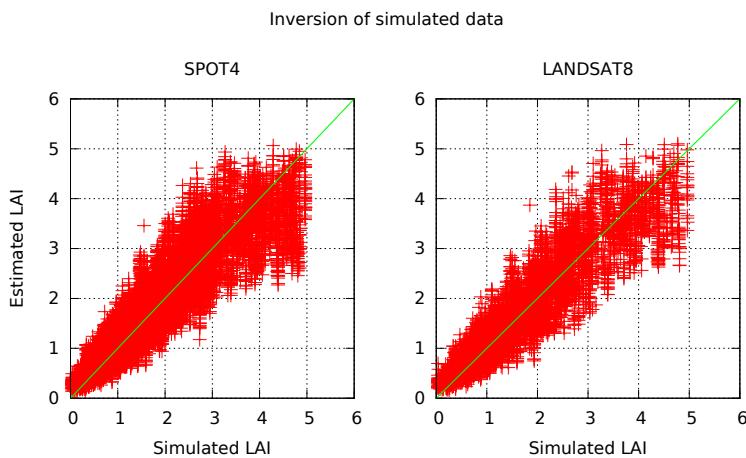


Figure 1: Scatter-plots for NLIRTM on simulated data

	SPOT	LANDSAT
Rel error(%)	22.19	22.40
RMSE	0.003	0.005

Table 1: Theoretical performances of the NLIRTM algorithm

**2.2.1.2 Comparison with BV-NET** The CESBIO team has received the support from INRA (M. Weiss and F. Baret) for the validation of the algorithm and cross-comparisons of results between the original BV-NET and the new implementation have been carried out.

Figure 2 shows an example of comparison between INRA's implementation and CESBIO's one using field measurements performed with hemispherical photographs and processed with the CanEye software. The scatter-plot shows similar performances for both approaches. Tables 2 and 3 show the relative errors and the RMSE for the validation points and allow to see that CESBIO's implementation yields slightly worse results. These

differences are still under investigation, but they seem to be produced by the different parameters used in PROSPECT as well as by the different methods to introduce soil reflectances in SAIL.

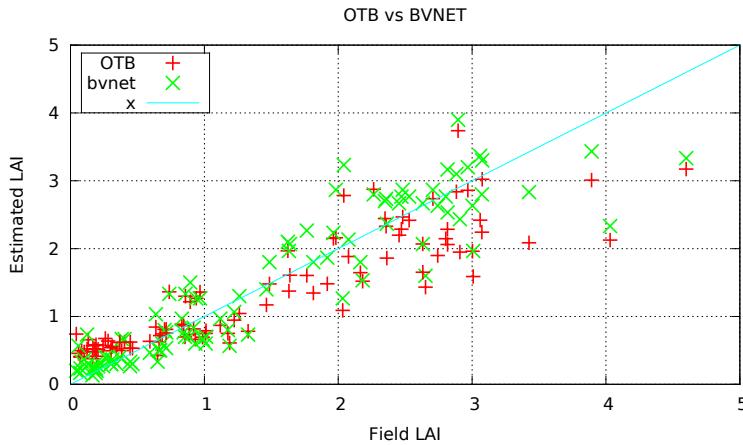


Figure 2: Scatter-plots for NLIRTM with *in situ* interpolated data

	SPOT	LANDSAT
BVNET	22.36	22.67
OTB	27.90	24.60

Table 2: Relative errors in %

	SPOT	LANDSAT
BVNET	0.15	0.04
OTB	0.19	0.06

Table 3: RMSE values

### 2.2.2 Algorithm setup

All results presented in this report use the surface reflectances, the NDVI and the RVI (Ratio Vegetation Index, that is the NIR reflectance divided by the Red one) as predictor variables. Instead of using the sensor and solar angles as predictors, a regression model is built per acquisition geometry.

The machine learning algorithm used for the inversion of the model is the Support Vector Regression since it needs fewer training samples and is parameter free. This is an advantage with respect to the Neural Network regression, which needs a high number of samples (which have to be simulated for each viewing configuration) and the inversion results are very sensitive to the architecture of the network (number of layers and number of neurons per layer). Support Vector Regression yields results of the same quality as Neural Networks at a lower cost.

## 2.3 LAI on-line multi-temporal retrieval (OLMTR)

The high temporal resolution of Sentinel-2 will allow the implementation of multi-temporal approaches for the LAI retrieval.

The approach proposed here builds upon the NLIRTM algorithm described in section 2.2 by using the LAI retrieval for several dates and combining them to improve the estimation.

The reprocessing of acquisition  $n$  will use all the acquisitions in a temporal window to perform a linear combination weighted by the temporal distance and the error estimation provided by the NLIRTM algorithm :

$$LAI_{OLMTR}(t_0) = \sum_{t_i=t_{first}}^{t_{last}} w_i LAI_{NLIRTM}(t_i). \quad (3)$$

The weights  $w_i$  are computed as follows:

$$w_i = \alpha_t \frac{1}{1 + |t_i - t_0|} + \alpha_\epsilon \frac{1}{1 + \epsilon_{NLIRTM}(t_i)}, \quad (4)$$

where  $\alpha_t$  and  $\alpha_\epsilon$  are the relative weights given to the temporal distance and the LAI retrieval errors.

For the present document, a time window of 3 acquisitions is used and the relative weights  $\alpha_t$  and  $\alpha_\epsilon$  are both set to 1. Two configurations are used:

- a symmetric window for the reprocessing of the penultimate acquisition;
- an asymmetric window using the last 3 acquisitions to reprocess the last one.

These 2 configurations will be called "N1" and "BW" (for backward window) respectively.

## 2.4 Complete season LAI multi-temporal retrieval (CSMTR)

This algorithm performs a reprocessing of the LAI retrievals of the complete season by fitting a double logistic function. The algorithm uses CESBIO's open source library PHENOTB<sup>1</sup> which implements a two-step estimation which is robust to the presence of several vegetative cycles in the season. Only the main cycle is fitted and the double logistic parameters are estimated. See section 2.5.1 for more details on the double logistic function.

Figure 3 illustrates some examples of double logistic fitting on real LAI profiles. One can observe that the algorithm is robust to noise and secondary cycles.

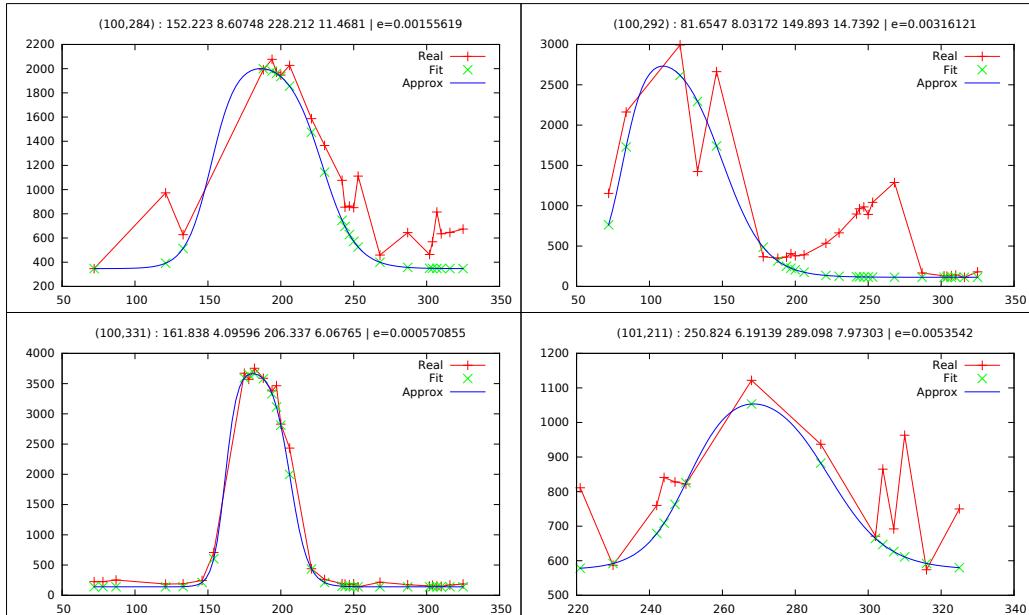


Figure 3: Examples of double logistic fitting on real LAI time series

This algorithm will be applied to the output of the NLIRTM algorithm at the end of the season. It will be referred as "LAI Fit" in the remainder of this document.

<sup>1</sup><http://tully.ups-tlse.fr/jordi/phenotb>

## 2.5 Temporal metrics from the NDVI temporal profile

The following phenological parameters were selected as meaningful to describe the vegetation status: emergence date, date of the maximum growth, length of the maturity plateau and senescence slope. Although the senescence slope was not included in the PSD selection, it has been investigated here as an additional useful metric.

These metrics can be extracted from time profiles of vegetation indices or biophysical variables yielding the same type of information. The use of vegetation indices seems more appropriate since the errors on the retrieval of the biophysical parameters are avoided.

### 2.5.1 Definitions

The logistic function has the form:

$$f(x) = \frac{1}{1 + e^{\frac{x_0 - x}{x_1}}} \quad (5)$$

The double logistic is:

$$g(x) = A(f_1(x) - f_2(x)) + B = A\left(\frac{1}{1 + e^{\frac{x_0 - x}{x_1}}} - \frac{1}{1 + e^{\frac{x_2 - x}{x_3}}}\right) + B, \quad (6)$$

where  $A + B$  is the maximum value and  $B$  is the minimum.

Since

$$\frac{df(x)}{dx} = \frac{e^{\frac{x_0 - x}{x_1}}}{x_1 \left(1 + e^{\frac{x_0 - x}{x_1}}\right)^2} \quad (7)$$

we have

$$g'(x) = \frac{dg(x)}{dx} = A\left(\frac{e^{\frac{x_0 - x}{x_1}}}{x_1 \left(1 + e^{\frac{x_0 - x}{x_1}}\right)^2} - \frac{e^{\frac{x_2 - x}{x_3}}}{x_3 \left(1 + e^{\frac{x_2 - x}{x_3}}\right)^2}\right) \quad (8)$$

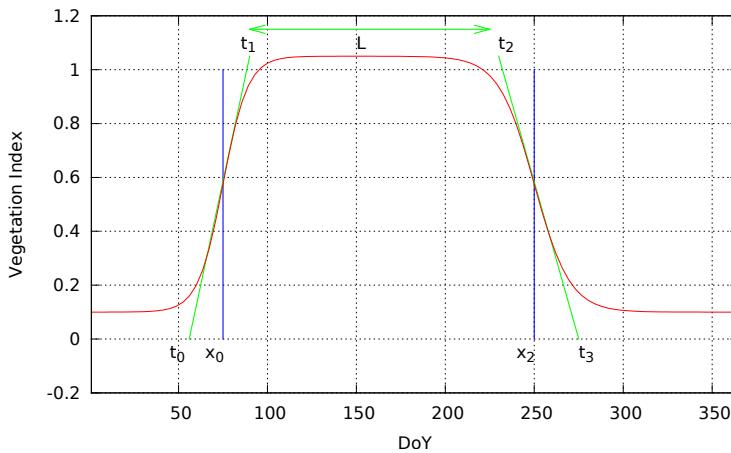


Figure 4: Double logistic function and associated parameters

**2.5.1.1 Date of the maximum positive gradient** By definition, this is  $x_0$ .

**2.5.1.2 Starting date** The date for which the straight line with the slope of  $x_0$  intercepts the horizontal axis. If

$$g'(x_0) = m \quad (9)$$

this line's equation is

$$y = mx + b \quad (10)$$

and verifies that

$$\begin{aligned} 0 &= mt_0 + b \\ g(x_0) &= mx_0 + b \end{aligned} \quad (11)$$

which gives

$$y = g'(x_0)x + (g(x_0) - g'(x_0)x_0) \quad (12)$$

and therefore,

$$t_0 = \frac{mx_0 - g(x_0)}{m} = x_0 - \frac{g(x_0)}{g'(x_0)} \quad (13)$$

**2.5.1.3 Length of the plateau** We define  $t_1$  as the date for which the previous straight line reaches the maximum value.

$$t_1 = \frac{A + B - (g(x_0) - g'(x_0)x_0)}{g'(x_0)} \quad (14)$$

Similarly, we can use the straight line associated to the descending slope:

$$y = g'(x_2)x + (g(x_2) - g'(x_2)x_2) \quad (15)$$

and define

$$t_2 = \frac{A + B - (g(x_2) - g'(x_2)x_2)}{g'(x_2)} \quad (16)$$

And the length of the plateau is:

$$L = t_2 - t_1 \quad (17)$$

**2.5.1.4 Senescence slope** By definition this is  $g'(x_2)$ .

## 2.5.2 Estimation

The metrics defined above are estimated on temporal NDVI profiles by using the same algorithm used for the CSMTR (see section 2.4). The output of the algorithm are the 4 double logistic parameters  $x_0, x_1, x_2, x_3$  and the minimum  $B$  and amplitude  $A$ .

The metrics are then obtained by applying the analytical expressions defined above.

## 2.5.3 Attempts to estimate the emergence date in real time

During the PDR held at CESBIO in February 2015 it was suggested to investigate the possibility of providing an estimation of the emergence date earlier than the end of the season.

Most of the methods in the literature estimate phenological parameters using either sigmoidal fitting based on the work of Zhang et al. [ZFS<sup>03</sup>] or empirical thresholds on vegetation indices. The most frequently estimated parameter is the onset of the vegetation, although it has many different definitions and names (SOS for start of season or start of spring, emergence date, etc.)

Many papers have compared different approaches on different settings: forests, grasslands and shrublands [BAH<sup>06</sup>]; croplands and forests [SH09]; different types of vegetation [CPC<sup>12</sup>], [TMW<sup>08</sup>]; crops [Tha12]; forests [WPS14].

All these papers use the full vegetation cycle either for fitting a parametric model or in order to produce a smoothed profile on which robust estimators can be applied. The key dates are either estimated from the parameters of the fitted models, or by using constraints on the derivatives of the profiles. For instance, Zangh et al. [ZFS<sup>03</sup>] propose the first maximum of the 3rd derivative of the double logistic function, but other versions using the 2nd derivative (the camel-back method [BGG<sup>07</sup>]) also exist.

The simpler methods using empirical thresholds on the VI profiles [RBV<sup>94</sup>], [MKVD97], need both a smoothing window of several months and the knowledge of the maximum value of the VI. For instance, Reed et al. use 9 14-day composites, which corresponds to more than 4 months. Furthermore, the thresholds are dependent on the vegetation type and the site.

All the literature cited above uses medium or low resolution imagery (AVHRR and MODIS) which ensures regular cloud-free data and smooth temporal behaviors. Only recently, have these approaches been compared using high spatial resolution data (30 m) [PHZ<sup>15</sup>], and even though a 2-day revisit cycle was used, the whole season approach was adopted.

These findings in the literature have been confirmed at CESBIO running some of these approaches on both simulated and Spot4 Take5 data. Some interesting perspectives have been identified, but more research has to be done in order to be able to propose a general algorithm for an operational processing chain, as required by the Sentinel2 Agriculture project.

As solution for the users to be able to estimate an emergence date earlier in the season, the NDVI for each acquisition could be provided, so that empirical thresholds specific to the site and the vegetation type can be computed. However, users able to implement their own algorithms for the estimation of the emergence date will be able to compute the NDVI from the level 2A products.

## 3 Quality criteria

### 3.1 *In situ* data preparation

LAI field measures are not always performed at the satellite acquisition dates. In order to use these *in situ* data for the validation of the algorithms, they have to be modified to match the image acquisition dates.

This is only possible if the *in situ* measures have been obtained close enough to the acquisition dates or if there exist measures after and before an image acquisition and therefore they can be interpolated.

Figure 5 shows several examples of *in situ* measures for the France - Sudmipy test site. It can be seen that for some cases, the main vegetative cycle is not complete. For other cases, the temporal gaps between the measures is large and therefore interpolation may not be appropriate.

Figure 6 shows similar examples for the Belgium test site. One can also observe that field measures may also contain errors, as for instance point number 10 where there is a decrease of LAI at the end of May before a very high peak.

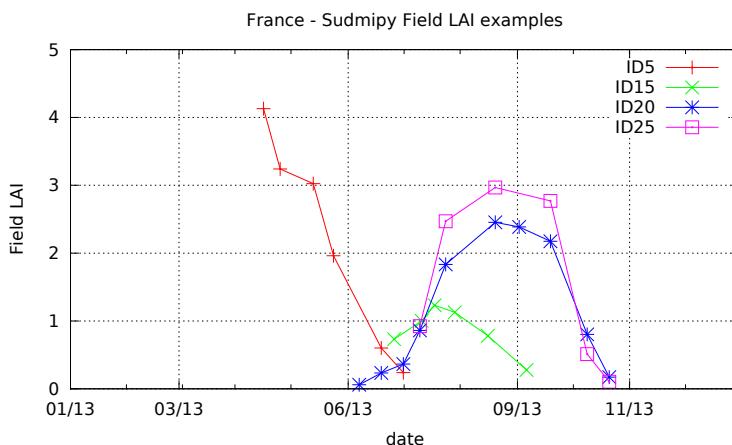


Figure 5: Examples of *in situ* data on France - Sudmipy

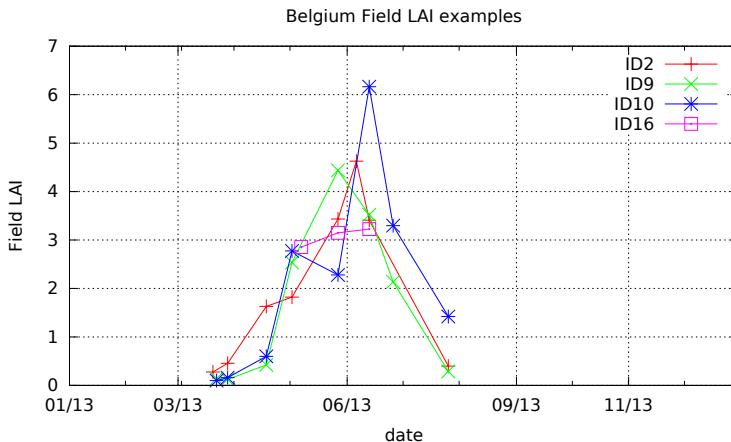


Figure 6: Examples of *in situ* data on Belgium

### 3.1.1 *In situ* data reliability issues

Field measures are supposed to be more reliable than satellite estimations and are therefore used here for validation. As stated above, the dates of the measures may not be appropriate for the satellite retrieval validation and some noise can also be present.

Two of the four test sites for which field LAI data were available, present also other types of issues.

The Morocco test site contains 10 field points (different geographical coordinates) which actually correspond to only 2 different measures (points 1 to 5 and points 6 to 10). The field values are averages of the real measures for the geographical points.

The Ukraine site presents the same set of fields measured over time but the coordinates of the points measured in each field are not consistent across the time and sometimes 2 points of the same field are measured at the same date but present large discrepancies in LAI values.

All these issues have led us to redefine the evaluation criteria proposed before the benchmarking phase.

## 3.2 Evaluation criteria

### 3.2.1 Initially proposed metrics

RMSE with respect to field measurements was proposed as the main evaluation criteria for the LAI retrieval. This needs the same dates for the satellite and the field measurements. Given the issues with field data listed above, a new error measure has been defined and is described below.

However, in order to have a reference error value for the benchmarking, *in situ* data interpolation was applied for the France - Sudmipy test site and RMSE and relative errors between these interpolated data and the satellite retrievals were computed for the NLIRTM algorithm.

The interpolation of the field data is performed as follows:

1. A temporal window of 7 days before and after each satellite acquisition is defined.
2. If field measures exist both before and after the satellite acquisition, a linear interpolation of the field measures is used as reference field value.
3. If only 1 field measure exists in the temporal window it is used as reference field value.
4. If no field measure is available in the temporal window, the satellite retrieval can not be validated and it is dropped.

It is worth noting that this approach leads to an important number of discarded field measures and this is the reason why only the France - Sudmipy test site is processed this way.

### 3.2.2 Time profile score for the LAI estimation

In order to exploit the available field data at its maximum potential, an error score which does not need interpolation has been designed. The main idea of the approach is to compute an error for each field measure using the available satellite retrievals in a temporal window. There is no interpolation of the satellite data in the score computation, because this would bias the score in favor of the retrieval algorithms which use some kind of temporal interpolation (the OLMTR and the CSMTR).

The error score is computed as follows:

1. For each field measure, a temporal window of N days before and after the current date is defined.
2. All satellite estimations in the temporal window are taken into account and contribute with an error score (equation 18).
3. The error score for a field measure is the mean of the score values in the temporal window.

The error score of each satellite retrieval has the form:

$$\frac{|LAI_{field} - LAI_{sat}|}{1 + (t_{field} - t_{sat})^2}, \quad (18)$$

that is, the difference in LAI weighted by the time gap between the field measure and the satellite retrieval.

This approach, used with a temporal window of 10 days before and after the field measure allows for the exploitation for all data in Belgium, France - Sudmipy and Morocco and about two thirds of the data in Ukraine.

Once the error score is evaluated for each field measure, an error score can be computed for the time profile by averaging the error scores of each date.

### 3.2.3 Criteria for the NDVI metrics

The RMSE for each of the selected metrics is used.

## 4 Benchmarking results

### 4.1 MLRISD

Two multi-linear regression models are estimated using the *in situ* data for all the sites, one model for SPOT4 Take5 data and another model for Landsat8 data.

Each model is evaluated using N-fold cross-validation. This technique consists in randomly splitting all the available data in N sets or folders, estimating the model using N-1 folder and validating it using the remaining folder. The procedure is repeated N times (one for each folder as validation set). Here, the number of folders is equal to the number of samples, which means that we estimate a regression model using all the samples but one and then we estimate the error using the remaining sample (leave one out validation). This is done for every sample in the test data set.

Since this algorithm needs *in situ* data for the model estimation, the validation is done only on the France - Sudmipy data set for which the *in situ* data has been interpolated to match some of the acquisition dates as described in section 3.2.1. This also allows to compare with the reference BV-NET implementation by INRA which was already used in section 2.2.1.

Figure 7 shows the scatter-plots of both the multi-linear regression and the BV-NET software for the case of SPOT4 Take5 and Landsat8 data. One can observe that the results are equivalent, the MLR being slightly better for Landsat and the inverse for SPOT4. This impression is backed up by tables 4 and 5 which present respectively the relative errors in % of LAI and the RMSE.

More than validating the MLR approach, these results confirm the state of the art for mid- and low-resolution LAI retrieval, which has shown that the inversion of radiative transfer models can achieve similar results to empirical regressions without the need for costly field measures.

It is worth noting that RMSE values are smaller for Landsat images, mainly due to better estimations of high LAI values. This result may be due to several causes as for instance:

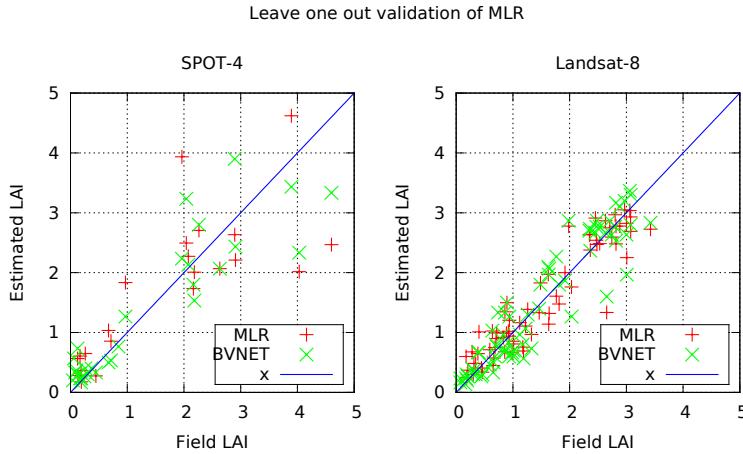


Figure 7: Leave one out validation of MLR and comparison with BV-NET

	SPOT	LANDSAT
BVNET	22.36	22.67
MLR	24.60	16.62

Table 4: Relative errors in %

- the period when the images were acquired (mostly in Summer for Landsat and in Spring for SPOT) and therefore the types of crops observed;
- the satellite viewing configuration closer to the nadir for Landsat producing measures more similar to the field hemispherical photographs.

The main difference with respect to lower resolutions is that here the errors are higher, since the geometry of the canopy (orientation of the rows of the crops in the fields, for instance) plays a role in high resolution imagery, but it is not taken into account by the models used in the regression.

## 4.2 NLIRTM

The results are already presented in figure 2 and tables 2 and 3 when compared with *is situ* interpolated data. Table 6 presents the error scores (see section 3.2.2 for the definition) obtained for each *in situ* site. These results don't give much information *per se*, but will serve as reference for the multi-temporal reprocessing.

Figure 8 presents some selected time profiles with the *in situ* data in red and the NLIRTM retrieval of the LAI in green. The error bars show the error estimation produced by the retrieval algorithm itself.

The complete time profiles are included in the appendix for reference.

## 4.3 Multitemporal estimations

Table 7 shows the summary of the scores (means and standard deviations) for the 4 test sites. The scores for each individual measure point are reported in the appendix with the plots for all the time profiles (starting on page 17).

The multi-temporal algorithms yield better results for the sites where the number of available images is higher (France - Sudmipy and Morocco). For the Belgium site, where the number of cloud-free acquisitions

	SPOT	LANDSAT
BVNET	0.15	0.04
MLR	0.21	0.04

Table 5: RMSE

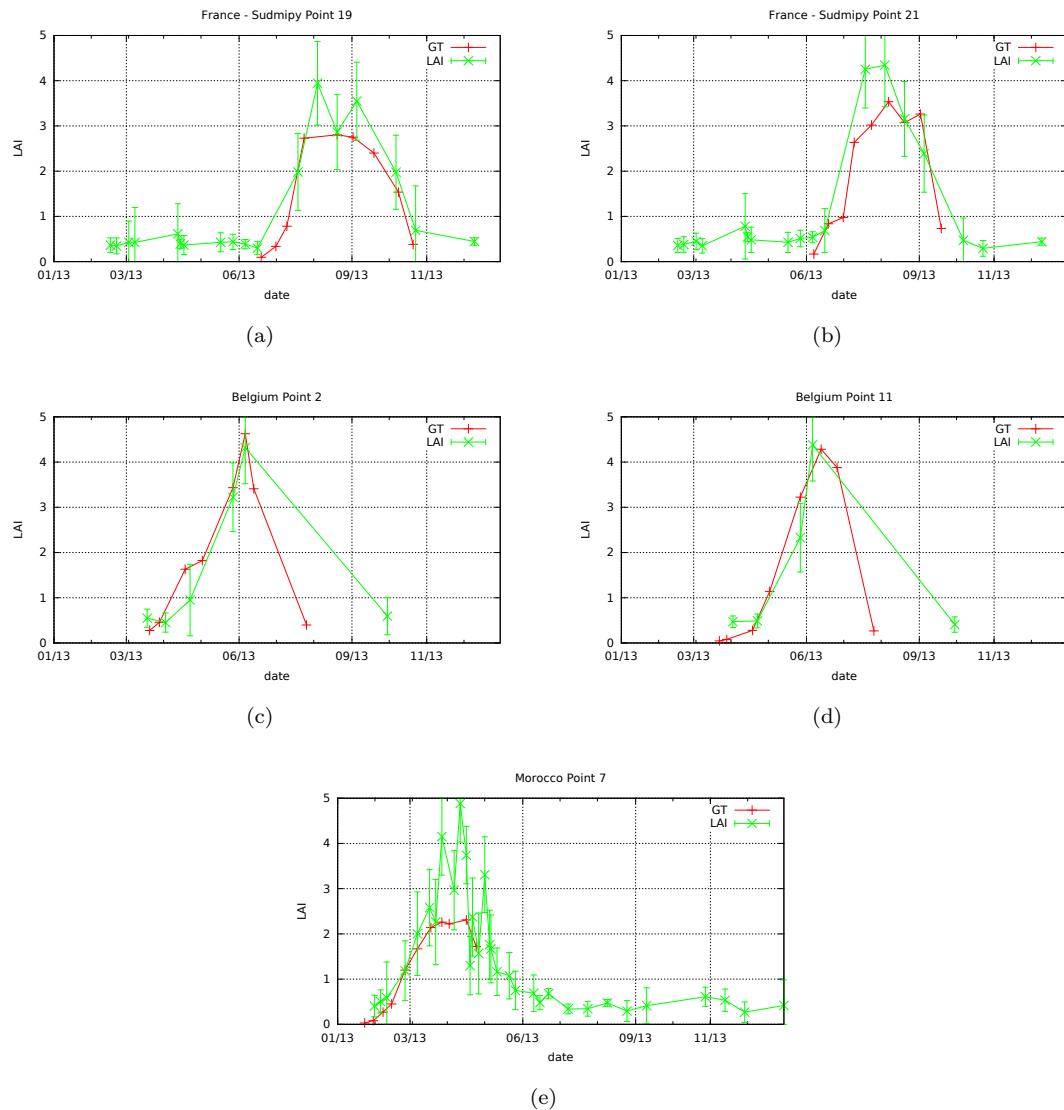


Figure 8: Examples of time profiles with the *in situ* data (Ground Truth) and the NLIRTM retrieval with the associated error estimations.

Site	mean	stdev
France - Sudmipy	0.15	0.23
Belgium	0.23	0.29
Morocco	0.25	0.15
Ukraine	0.10	0.26

Table 6: Error scores per site

site	NLIRTM		N1		BW		Fit	
	mean	stdev	mean	stdev	mean	stdev	mean	stdev
France - Sudmipy	0.15	0.23	<b>0.12</b>	0.16	<b>0.12</b>	0.17	0.13	0.20
Belgium	0.23	0.29	0.30	0.32	<b>0.17</b>	0.19	0.35	0.45
Morocco	0.25	0.15	0.17	0.12	0.20	0.14	<b>0.09</b>	0.06
Ukraine	0.10	0.26	<b>0.09</b>	0.21	0.11	0.23	0.10	0.23

Table 7: Error scores per site for the 3 algorithms. Best results for each site are in bold.

is very low, and the N1 algorithm yields worse results than the mono-date approach. However, the use of a backward window improves the estimation. Finally, the Ukraine site, where the reliability of the field measures seems bad, there is no clear winner algorithm.

#### 4.4 Temporal metrics

The validation of the temporal metrics need information which is not available in the test data set of the project. Indeed, the dates for the emergence and senescence and the growing rates of the crops were not recorded. It was therefore decided to simulate temporal profiles for which these parameters are known.

The drawback of this approach is that the quality of the estimations of the metrics may be overly optimistic with respect to what could be achieved on real data. In order to minimize this risk, the following simulation approach was adopted.

1. The 4 parameters  $x_0$  to  $x_3$  of the double logistic function are randomly selected as follows:
  - (a)  $x_0$  is drawn from a Gaussian distribution centered on day 90 with a standard deviation of 60 days.
  - (b)  $x_2$  is simulated as  $x_0$  plus a random Gaussian value with mean 70 and standard deviation 30, which allows for maturity plateaux spreading from a less than 10 days up to 100 days.
  - (c) Based on the analysis of many time profiles, the inverse rates of growth and senescence ( $x_1$  and  $x_3$ ) are generated using a Gaussian distribution with mean 7 and standard deviation 4.
2. The maximum possible value of the profile is set to 1 for NDVI and a random modulation of this value is performed in order to generate the profile amplitude  $A$  and the minimum value  $B$  as follows:
  - (a) the standard deviation  $\sigma$  is set to 1% of the maximum and used to draw values from a Gaussian distribution  $N(0, \sigma)$
  - (b)  $A = 1 - 3\sigma + N(0, \sigma)$
  - (c)  $B = 3\sigma + N(0, \sigma)$
3. A set of  $N$  acquisition dates is randomly generated using a uniform distribution on the interval [1, 365].
4. The double logistic function with the selected parameters is sampled at the  $N$  acquisition dates.
5. Random Gaussian noise of a selected variance is added to each profile value.

All the tests presented below use 20 dates in 12 months (which is about 30% of cloud-free images for a 5-day revisit cycle) and a noise standard deviation of 0.02. Some time profiles examples are illustrated in figure 9. As

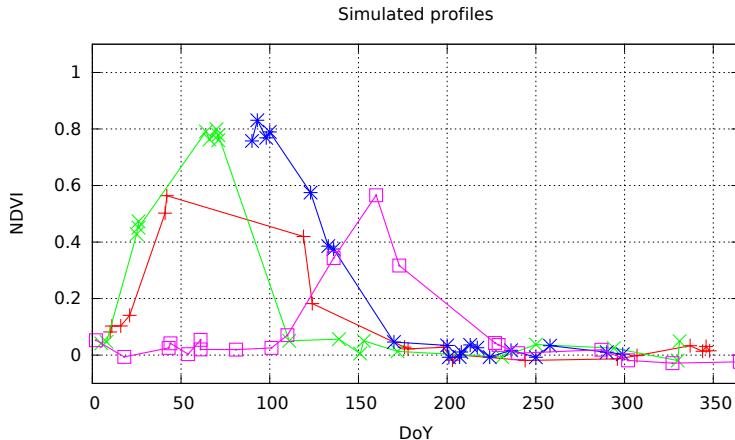


Figure 9: Examples of simulated NDVI time profiles

one can observe, these are quite realistic. For example a time profile can start late in the season (the blue one) or have long data gaps during the NDVI peak (the red one).

Using a set of 100 profiles, the scatter-plots of the estimations of the 4 metrics are present in figure 10. Table 8 reports the RMSE of the estimation of the 4 metrics.

The date of the maximum growth and the emergence date are correctly estimated within a 2 week range and the scatter-plots show good agreement with respect to the theoretical values.

Errors in the plateau length are important for very short (less than 20 days) and very long cycles (more than 80 days). Other than the fact that these seldom exist in reality, the errors come from either starting or ending dates at the limits of the season or too close to one another for the peak to be visible.

Finally, the senescence rate is not estimated reliably, since the derivative values are very sensitive to estimation noise. These results suggest that providing the senescence date ( $t_2$  in figure 4) would be more meaningful, since this parameter is estimated with the same reliability as  $t_0$ . The senescence rate could also be estimated as  $t_3 - t_2$  with a reliability similar to the one of the plateau length.

$x_0$	$t_0$	L	$dgx_2$
16.02	14.47	19.31	0.76

Table 8: RMSE values for the temporal metrics

## 5 Conclusions on algorithm selection

We propose to assess the vegetation status with 2 different products: an LAI retrieval throughout the season and a set of temporal metrics at the end of the season.

### 5.1 LAI retrieval

Several approaches for near-real-time LAI retrieval have been compared leading to the following conclusions:

1. Multi-linear regression using field data is an accurate approach, but not usable at a large scale because of the need of *in situ* data for model calibration. Furthermore, non linear regression for the inversion of PROSPECT+SAIL yields comparable results.
2. Mono-date estimations using the inversion of PROSPECT+SAIL produce LAI retrievals with errors less than 30%, which is in agreement with results found in the literature.
3. A sliding window for reprocessing a buffer of recent dates has been investigated to provide an improved version of the mono-date algorithm. A minor improvement is observed. However, only 2 configurations of the parameters of the algorithm have been explored.

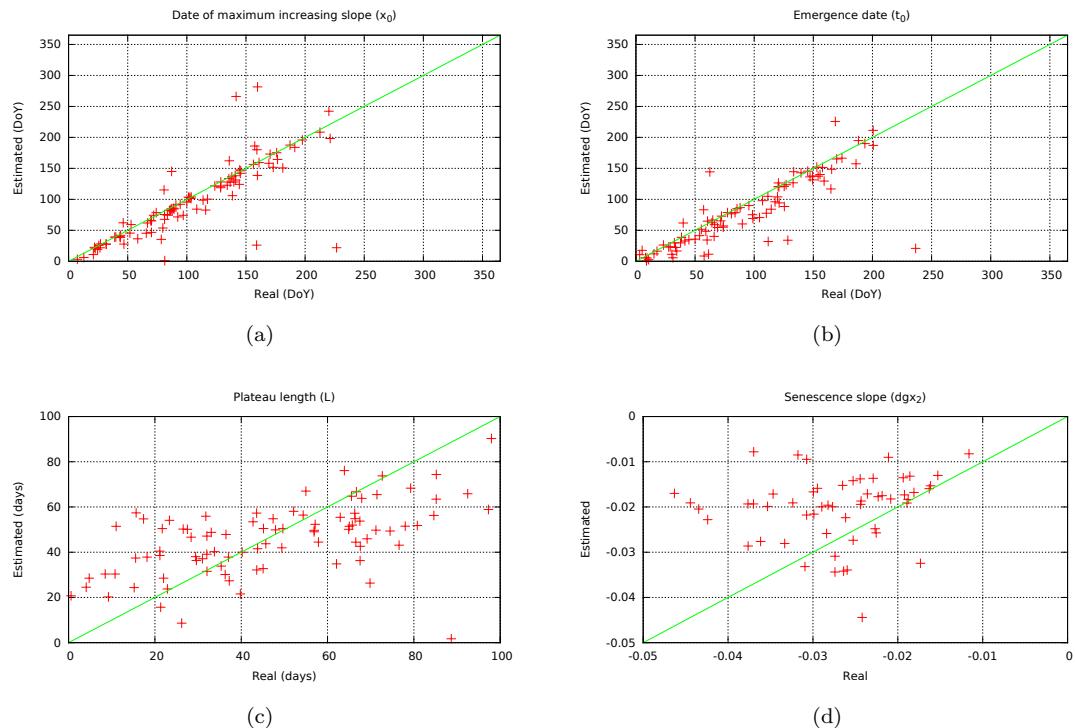


Figure 10: Scatter-plots of the estimation of the temporal metrics.

4. A phenological model fit at the end of the season improves the error scores for the LAI retrieval for the cases where a high number of cloud-free images is available.

As a consequence, the following processing system options are proposed:

1. Perform the non linear regression for every acquisition.
2. Keep open the option of performing a multi-temporal reprocessing and further investigate the algorithm parameters.
3. Reprocess the complete profiles at the end of the season with the fitting approach if the number of available acquisition exceeds a given threshold.

These choices allow to specify the algorithms so that the system can be developed. Further refinement on the algorithm parameters can still be done.

## 5.2 Temporal NDVI metrics

The results on simulated data suggest the following algorithm selection:

1. Provide emergence date, date of the maximum growth rate and length of the plateau as initially proposed.
2. Replace the senescence date initially proposed by the senescence date  $t_2$  or the senescence length defined as  $t_3 - t_2$ .

Since the OLMTR algorithm is also able to provide fAPAR estimates, it is proposed to investigate the possibility of performing the temporal metric estimation on this biophysical parameter. However, no *in situ* data is available for the validation within the Sentinel-2 Agriculture project.

## 6 References

- [BAH<sup>+</sup>06] Pieter S.A. Beck, Clement Atzberger, Kjell Arild Høgda, Bernt Johansen, and Andrew K. Skidmore. Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sensing of Environment*, 100(3):321–334, 2006.
- [BGG<sup>+</sup>07] Heiko Balzter, France Gerard, Charles George, Graham Weedon, Will Grey, Bruno Combal, Etienne Bartholomé, Sergey Bartalev, and Sietse Los. Coupling of vegetation growing season anomalies and fire activity with hemispheric and regional-scale climate patterns in Central and East Siberia. *Journal of Climate*, 20(15):3713–3729, 2007.
- [CPC<sup>+</sup>12] Nan Cong, Shilong Piao, Anping Chen, Xuhui Wang, Xin Lin, Shiping Chen, Shijie Han, Guangsheng Zhou, and Xinping Zhang. Spring vegetation green-up date in China inferred from SPOT NDVI data: A multiple model analysis. *Agricultural and Forest Meteorology*, 165(nil):104–113, 2012.
- [FFA<sup>+</sup>08] Jean-Baptiste Feret, Christophe François, Gregory P. Asner, Anatoly A. Gitelson, Roberta E. Martin, Luc P.R. Bidel, Susan L. Ustin, Guerric le Maire, and Stéphane Jacquemoud. Prospect-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment*, 112(6):3030–3043, 2008.
- [JVB<sup>+</sup>09] Stéphane Jacquemoud, Wout Verhoef, Frédéric Baret, Cédric Bacour, Pablo J. Zarco-Tejada, Gregory P. Asner, Christophe François, and Susan L. Ustin. Prospect+SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment*, 113(nil):S56–S66, 2009.
- [MKVD97] S. Moulin, L. Kergoat, N. Viovy, and G. Dedieu. Global-scale assessment of vegetation phenology using NOAA/AVHRR satellite measurements. *Journal of Climate*, 10(6):1154–1170, 1997.
- [PHZ<sup>+</sup>15] Zhuokun Pan, Jingfeng Huang, Qingbo Zhou, Limin Wang, Yongxiang Cheng, Hankui Zhang, George Alan Blackburn, Jing Yan, and Jianhong Liu. Mapping crop phenology using NDVI time-series derived from HJ-1 A/B data. *International Journal of Applied Earth Observation and Geoinformation*, 34(nil):188–197, 2015.
- [RBV<sup>+</sup>94] Bradley C. Reed, Jesslyn F. Brown, Darrel VanderZee, Thomas R. Loveland, James W. Merchant, and Donald O. Ohlen. Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, 5(5):703–714, 1994.
- [SH09] Mark D. Schwartz and Jonathan M. Hanes. Intercomparing multiple measures of the onset of spring in eastern north america. *Int. J. Climatol.*, 30(11):1614–1626, 2009.
- [Tha12] Jonathan B. Thayn. Assessing vegetation cover on the date of satellite-derived start of spring. *Remote Sensing Letters*, 3(8):721–728, 2012.
- [TMW<sup>+</sup>08] Bin Tan, Jeffrey T. Morissette, Robert E. Wolfe, Feng Gao, Gregory A. Ederer, Joanne Nightingale, and Jeffrey A. Pedelty. Vegetation phenology metrics derived from temporally smoothed and gap-filled MODIS data. In *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium*, page nil, nil 2008.
- [WBL<sup>+</sup>02] Marie Weiss, Frédéric Baret, Marc Leroy, Olivier Hautecœur, Cédric Bacour, Laurent Prévot, and Nadine Bruguière. Validation of neural net techniques to estimate canopy biophysical variables from remote sensing data. *Agronomie*, 22(6):547–553, 2002.
- [WPS14] Katharine White, Jennifer Pontius, and Paul Schaberg. Remote sensing of spring phenology in northeastern forests: A comparison of methods, field metrics and sources of uncertainty. *Remote Sensing of Environment*, 148(nil):97–107, 2014.

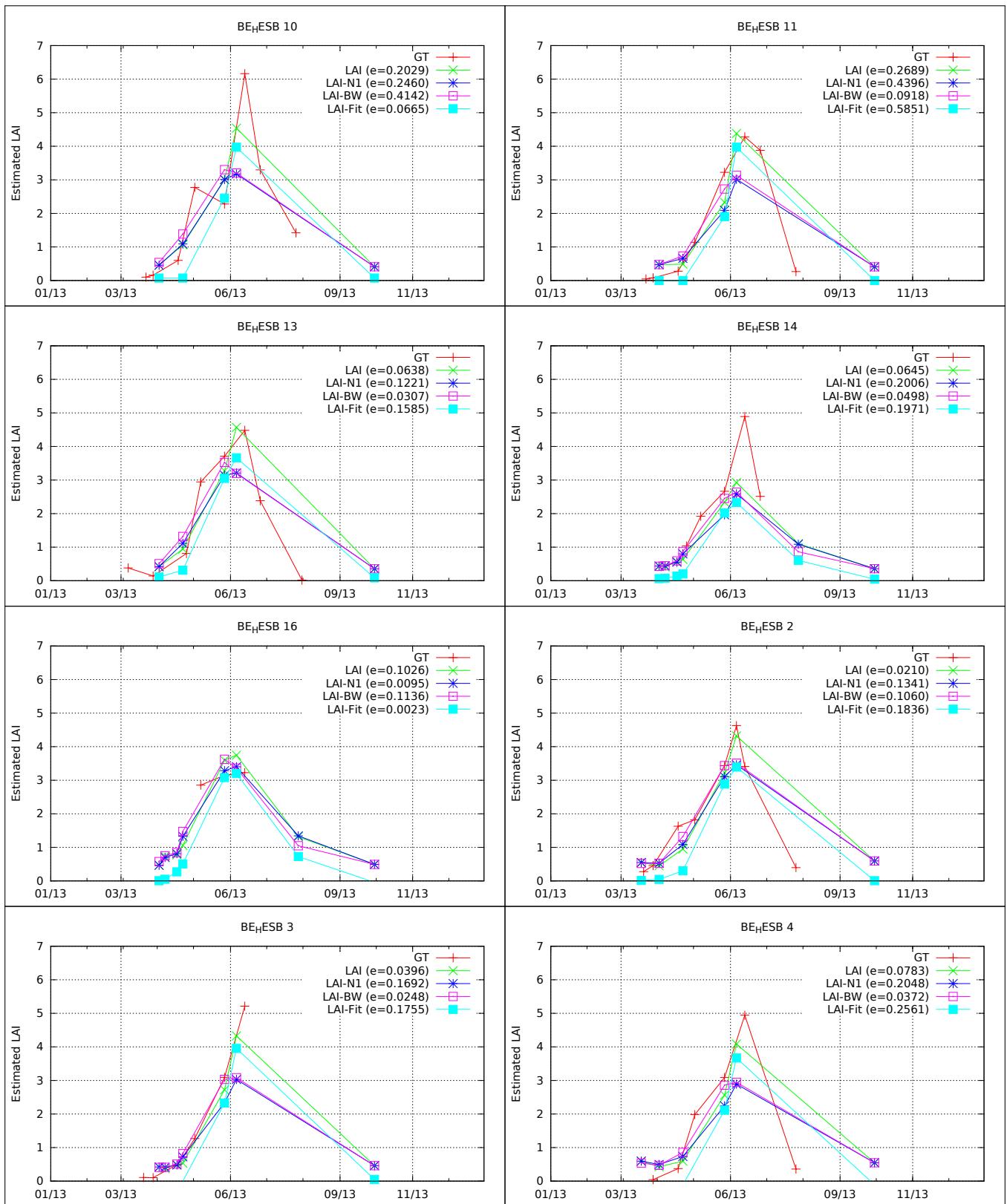
- 
- [ZFS<sup>+03</sup>] Xiaoyang Zhang, Mark A. Friedl, Crystal B. Schaaf, Alan H. Strahler, John C.F. Hodges, Feng Gao, Bradley C. Reed, and Alfredo Huete. Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, 84(3):471–475, 2003.

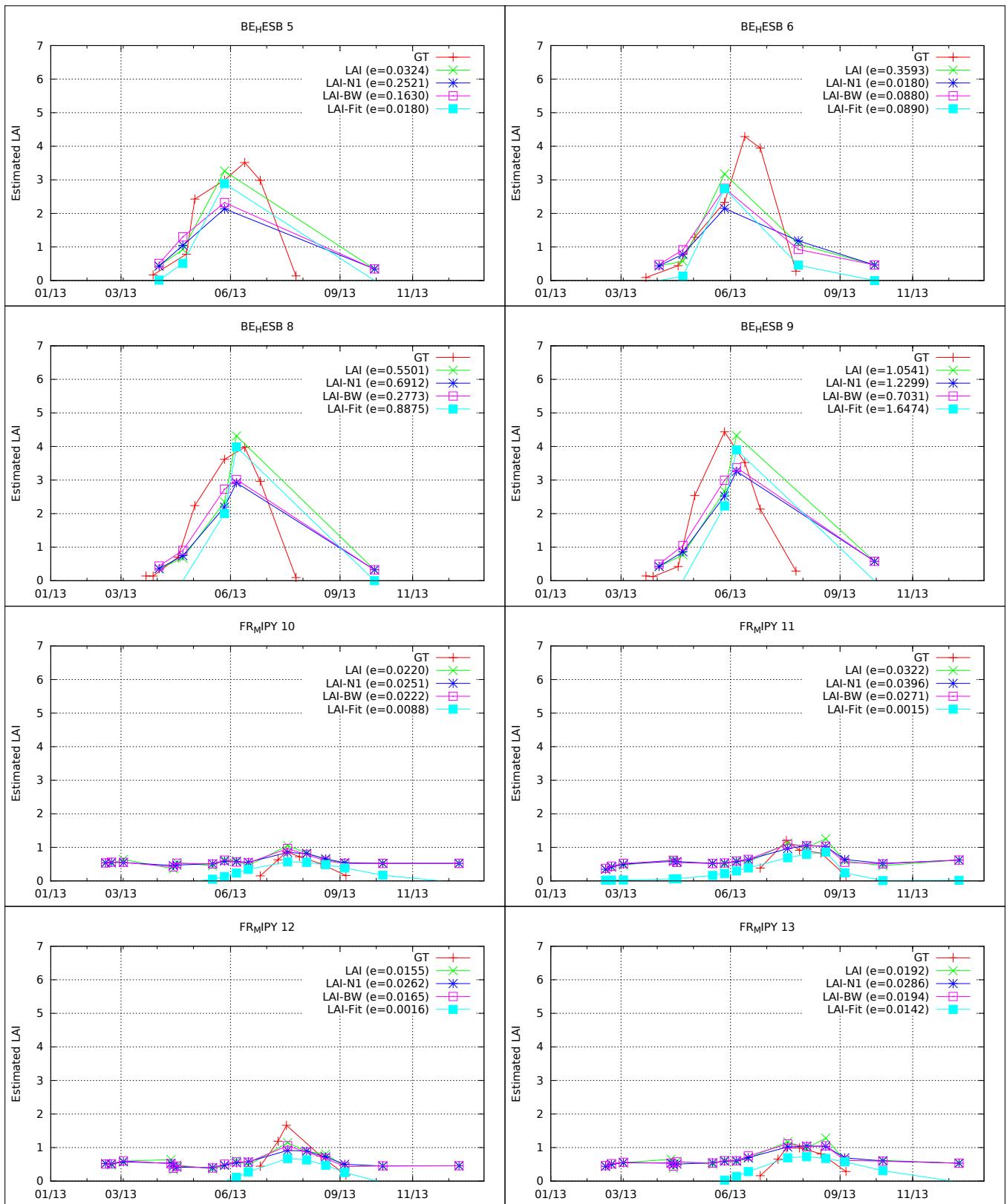
## 7 Appendix

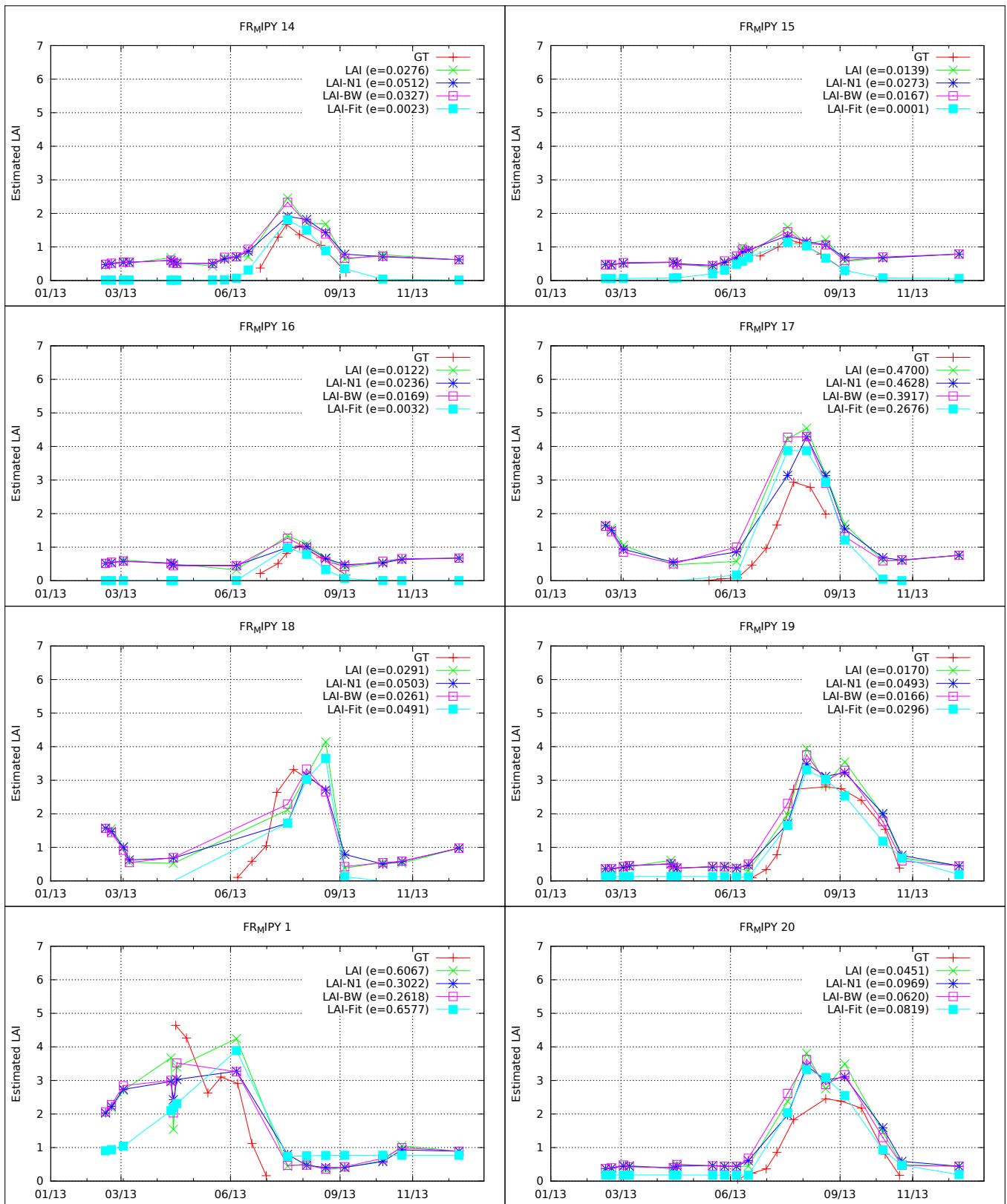
### 7.1 Temporal profiles

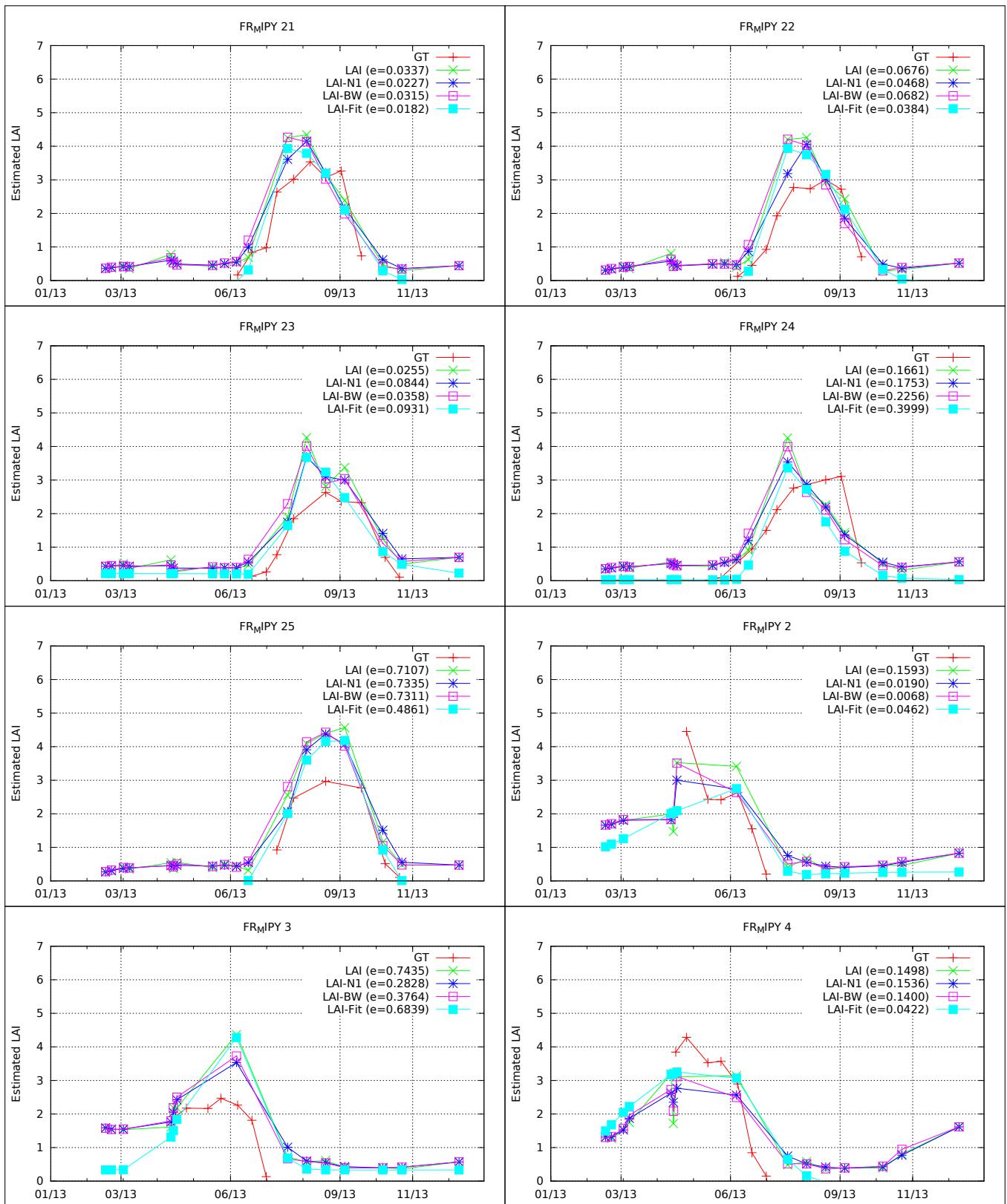
The following pages present the complete set of temporal profiles processed with the NLIRTM, OLMTR and CSMTR algorithms. The field measures appear in red (GT for Ground Truth), the NLIRTM appears in green (LAI), the OLMTR (LAI-N1in blue, LAI-BW in pink) and the CSMTR in cyan (LAI-Fit). Each legend reports the average score presented in section 3.2.2.

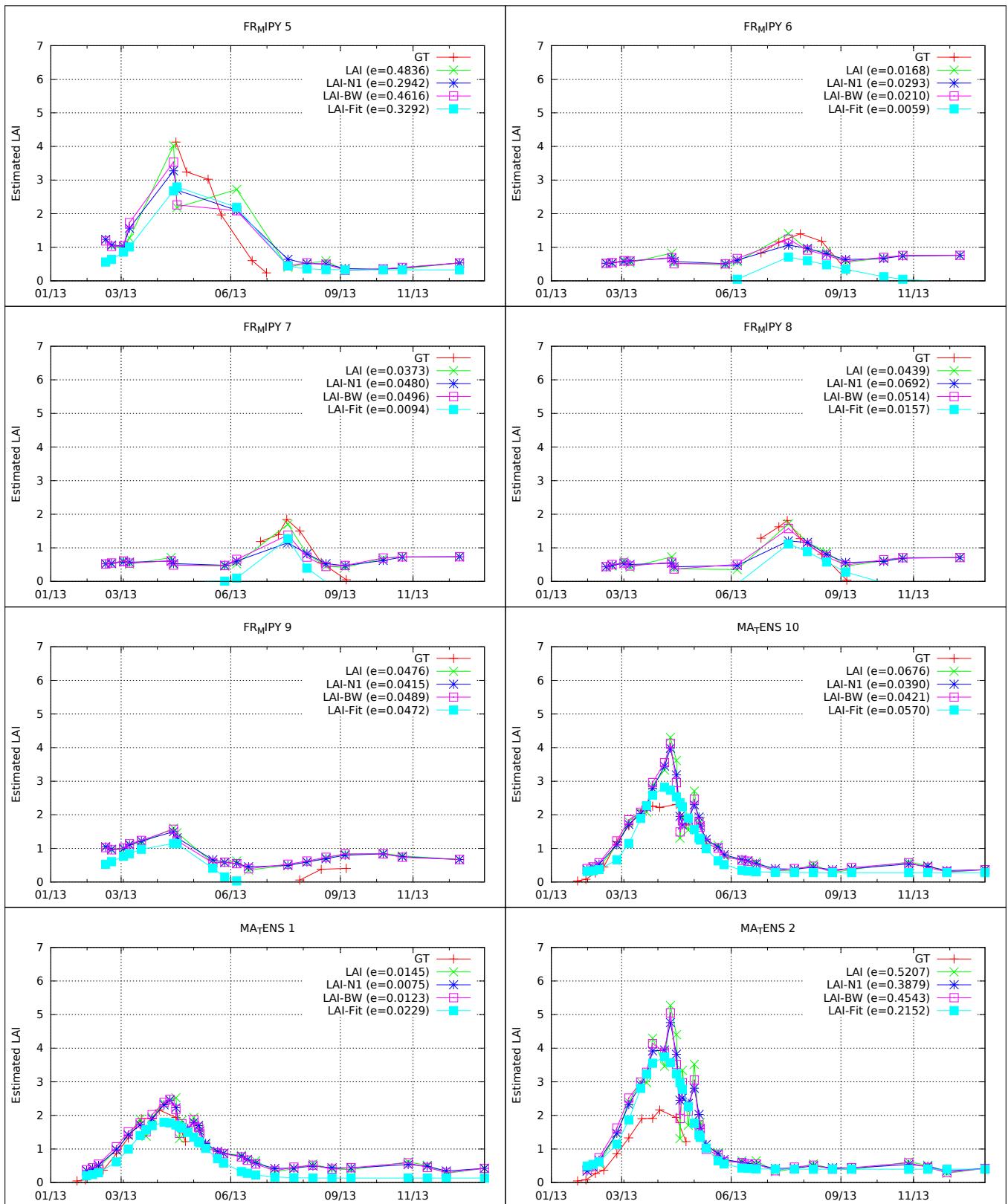
For the case of the Ukraine test site, the scores are computed with respect one single date each time, for the reasons explained in section 3.1.1. In some cases, all satellite measures are more than 10 days away from the field measure and the score can not be computed. It is therefore reported as *nil*.

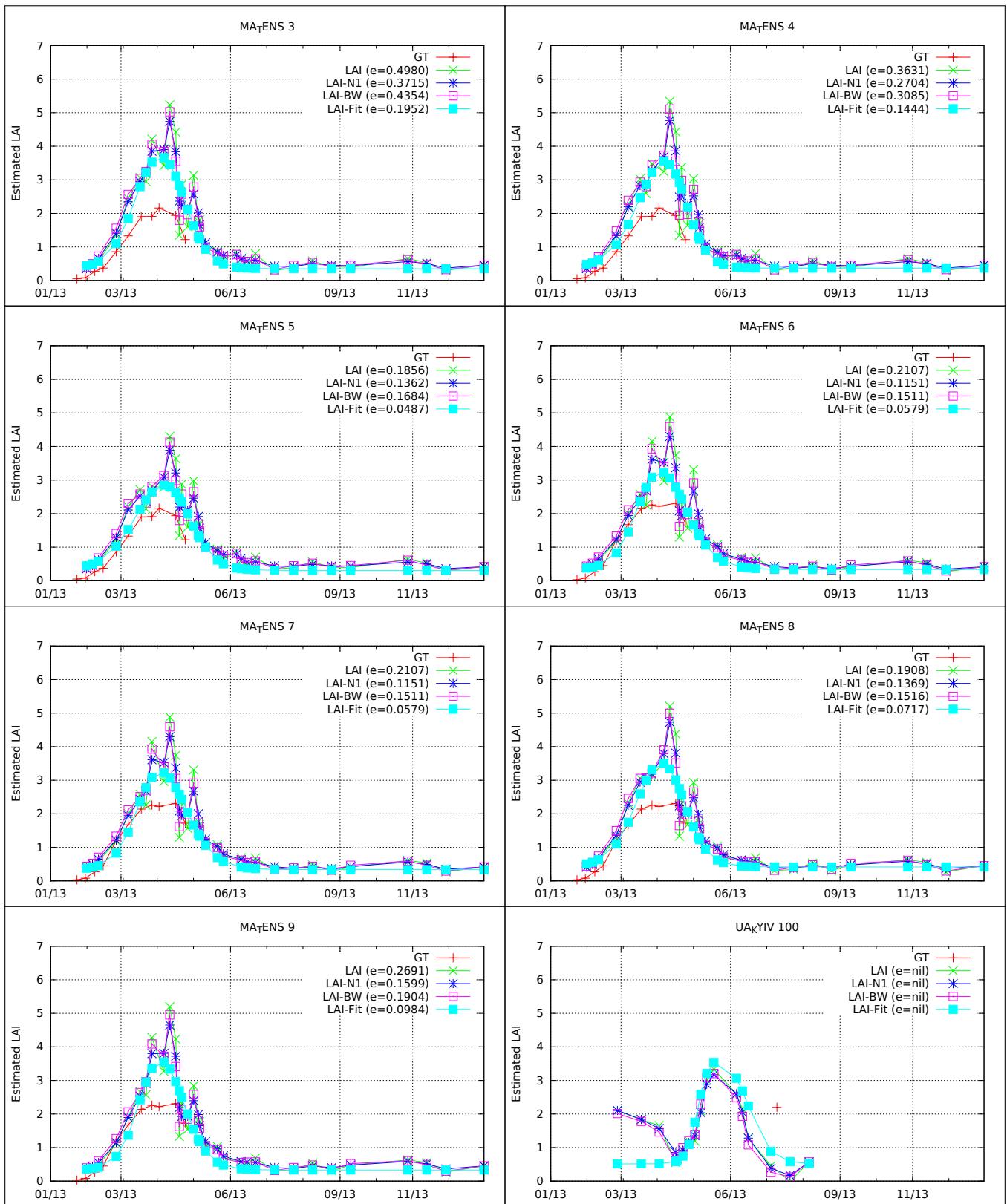


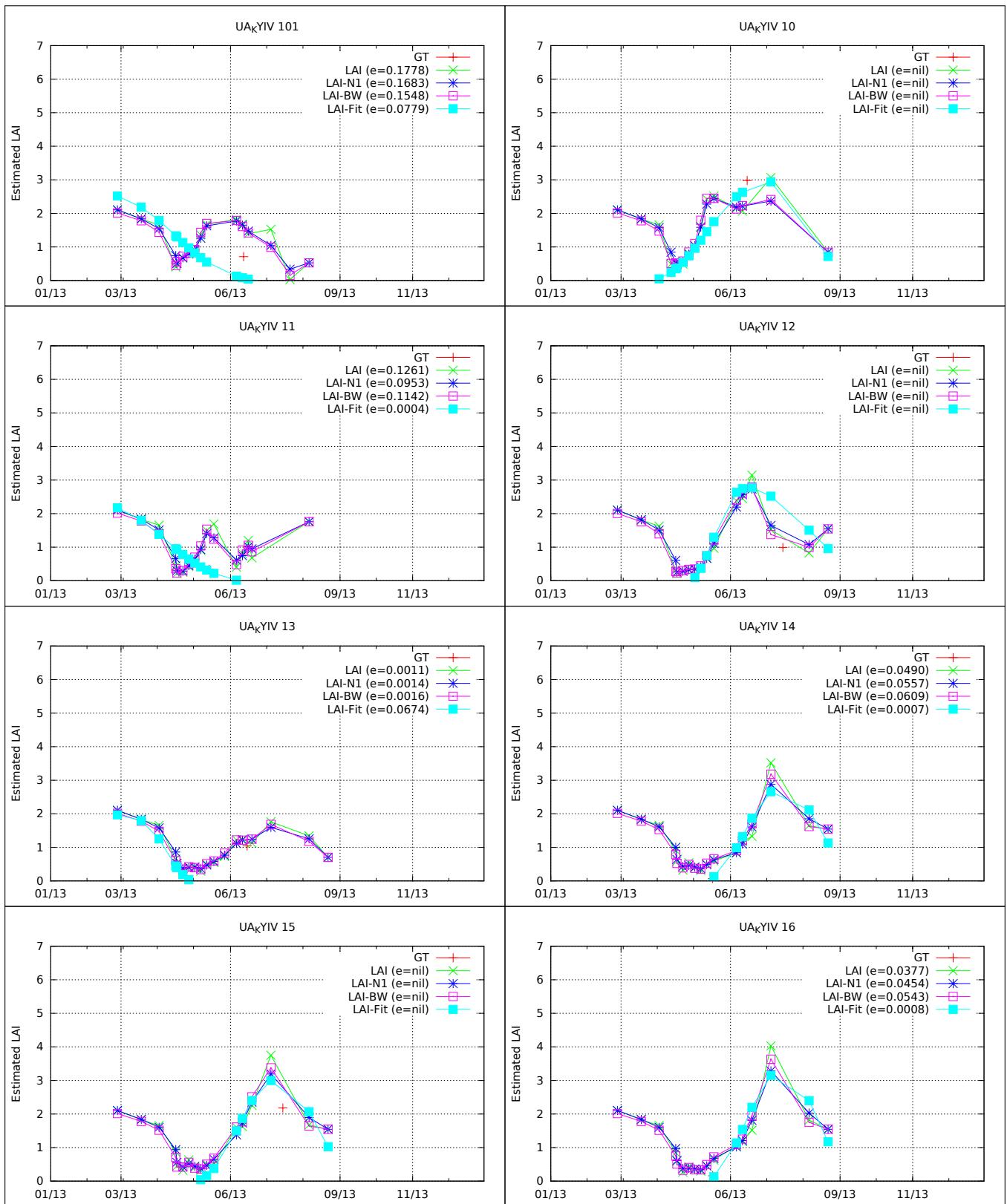


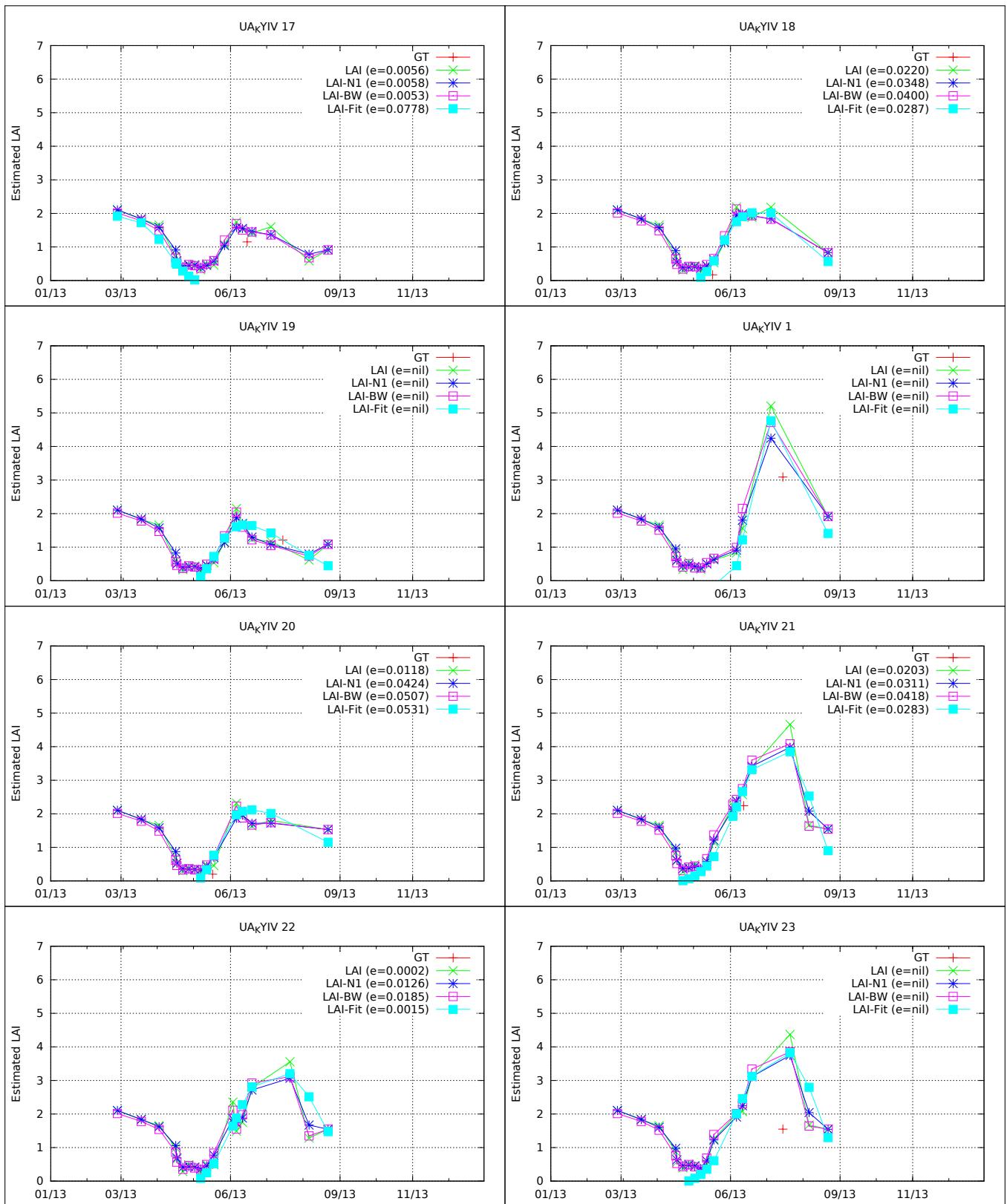


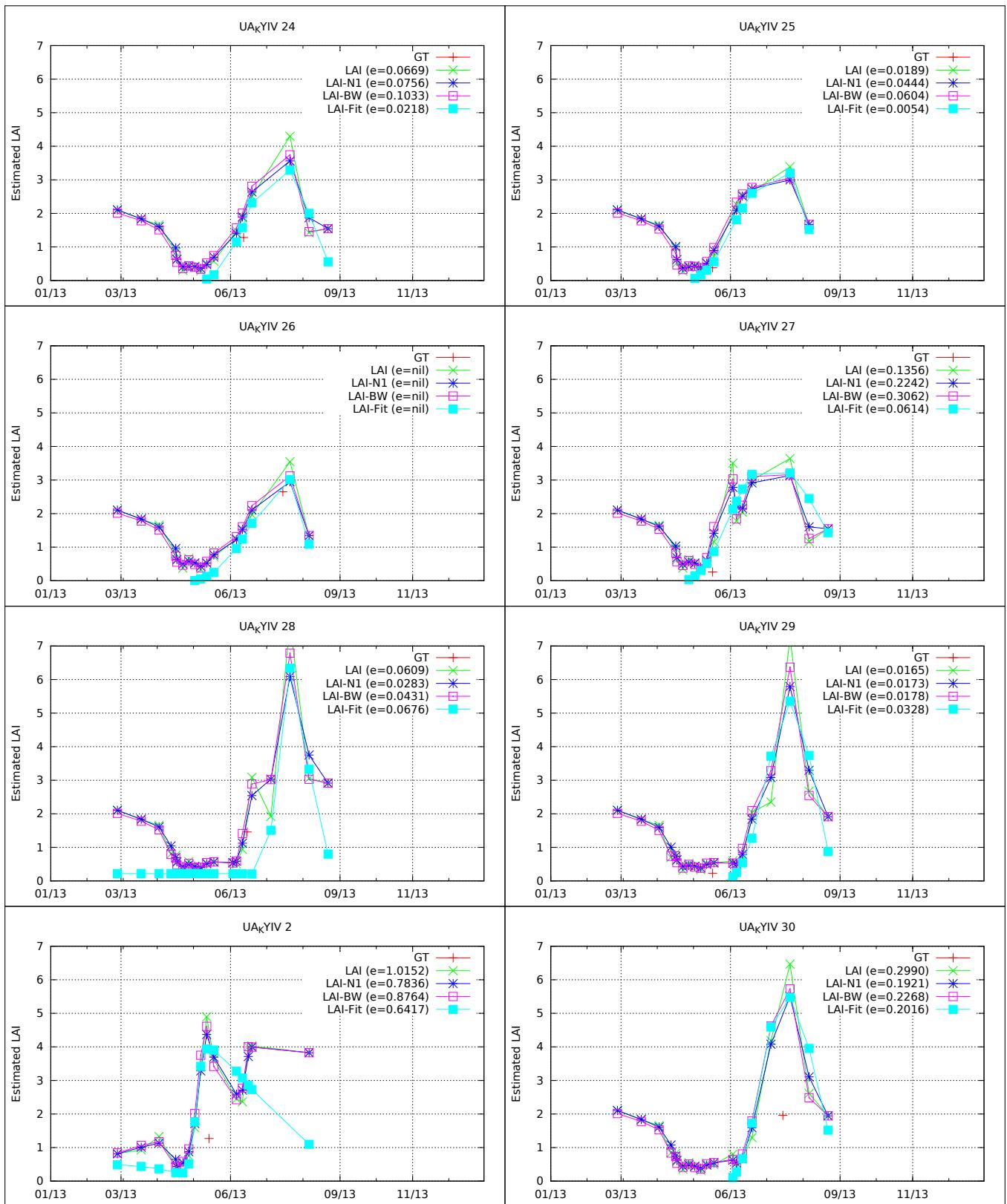


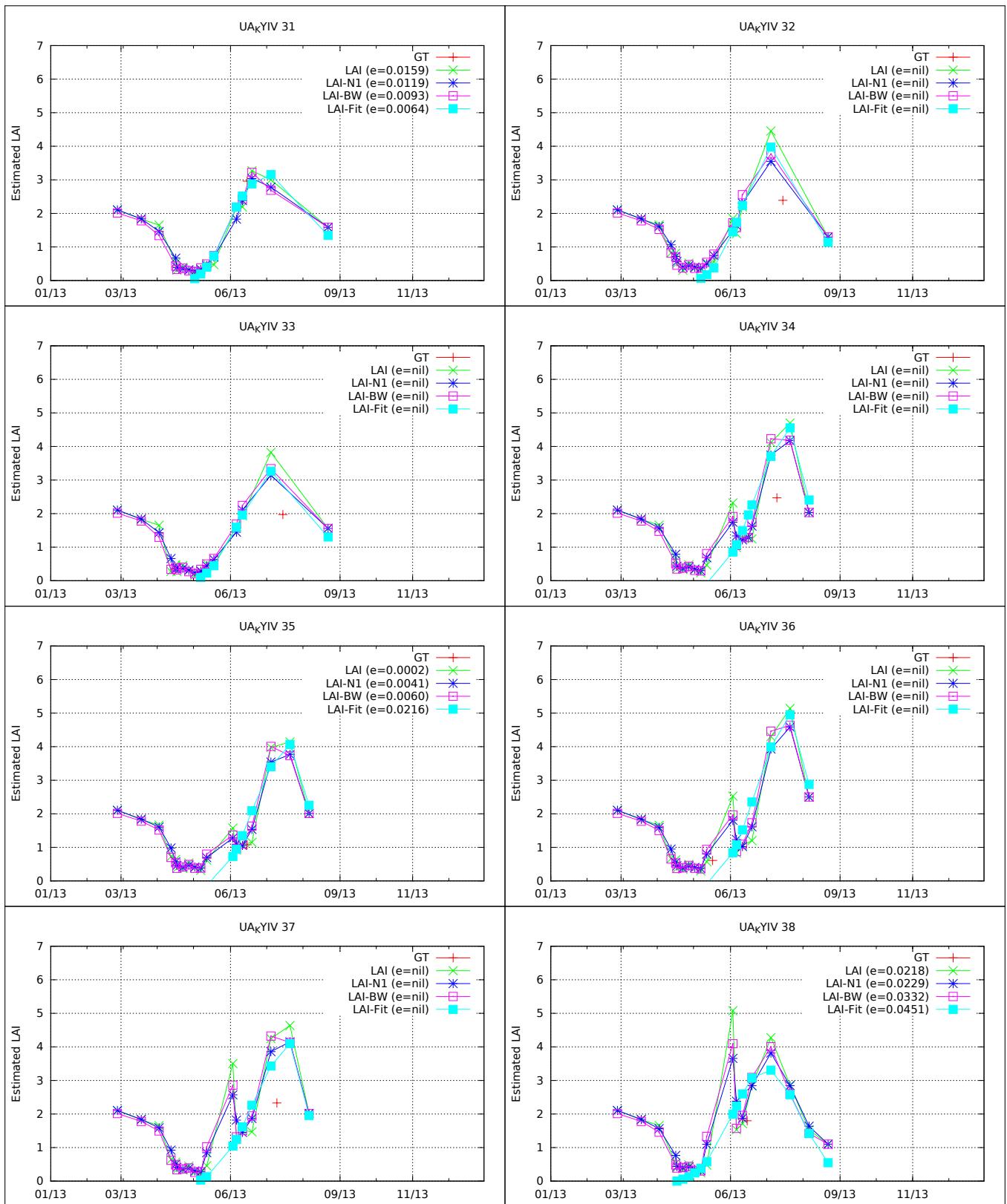


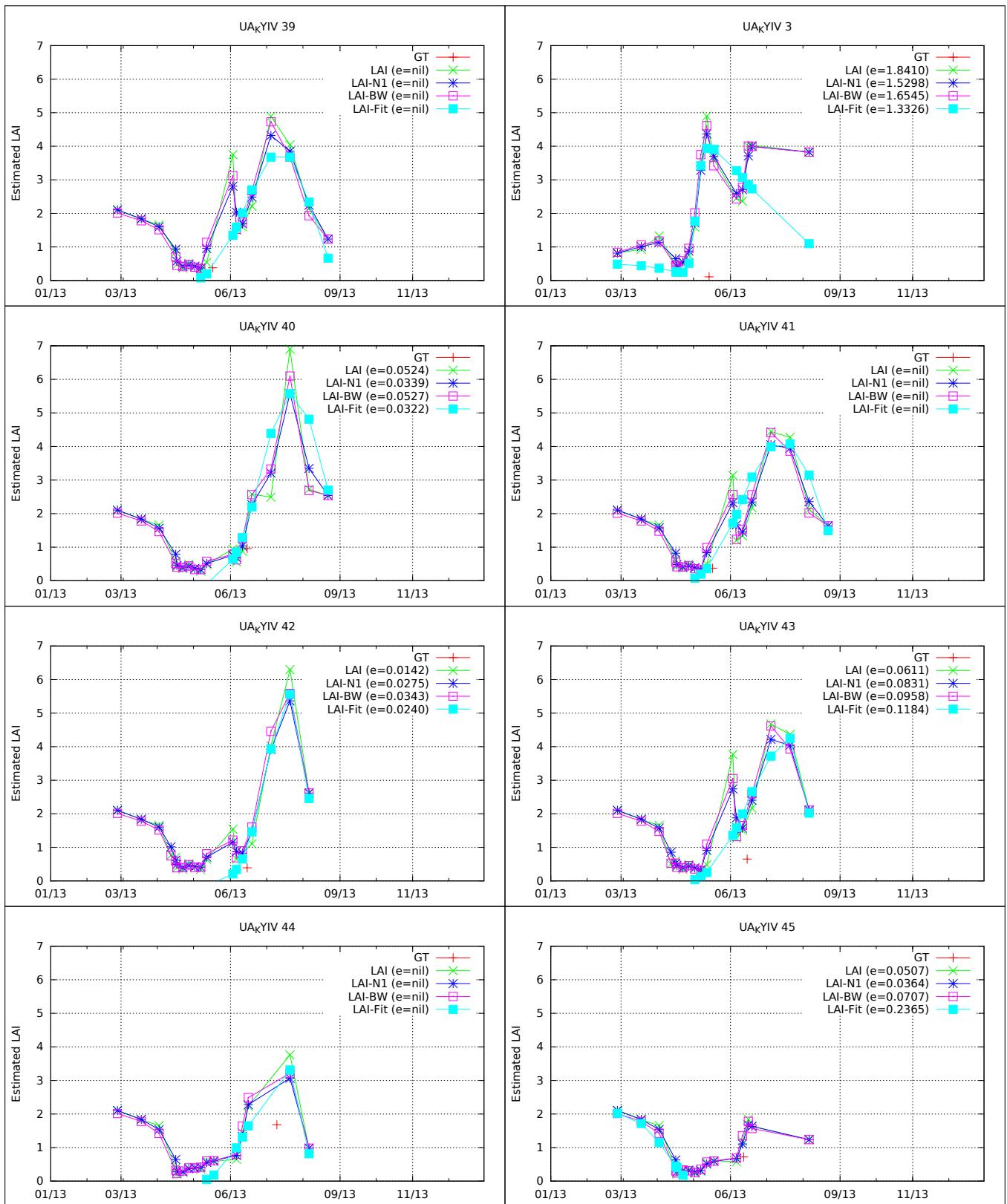


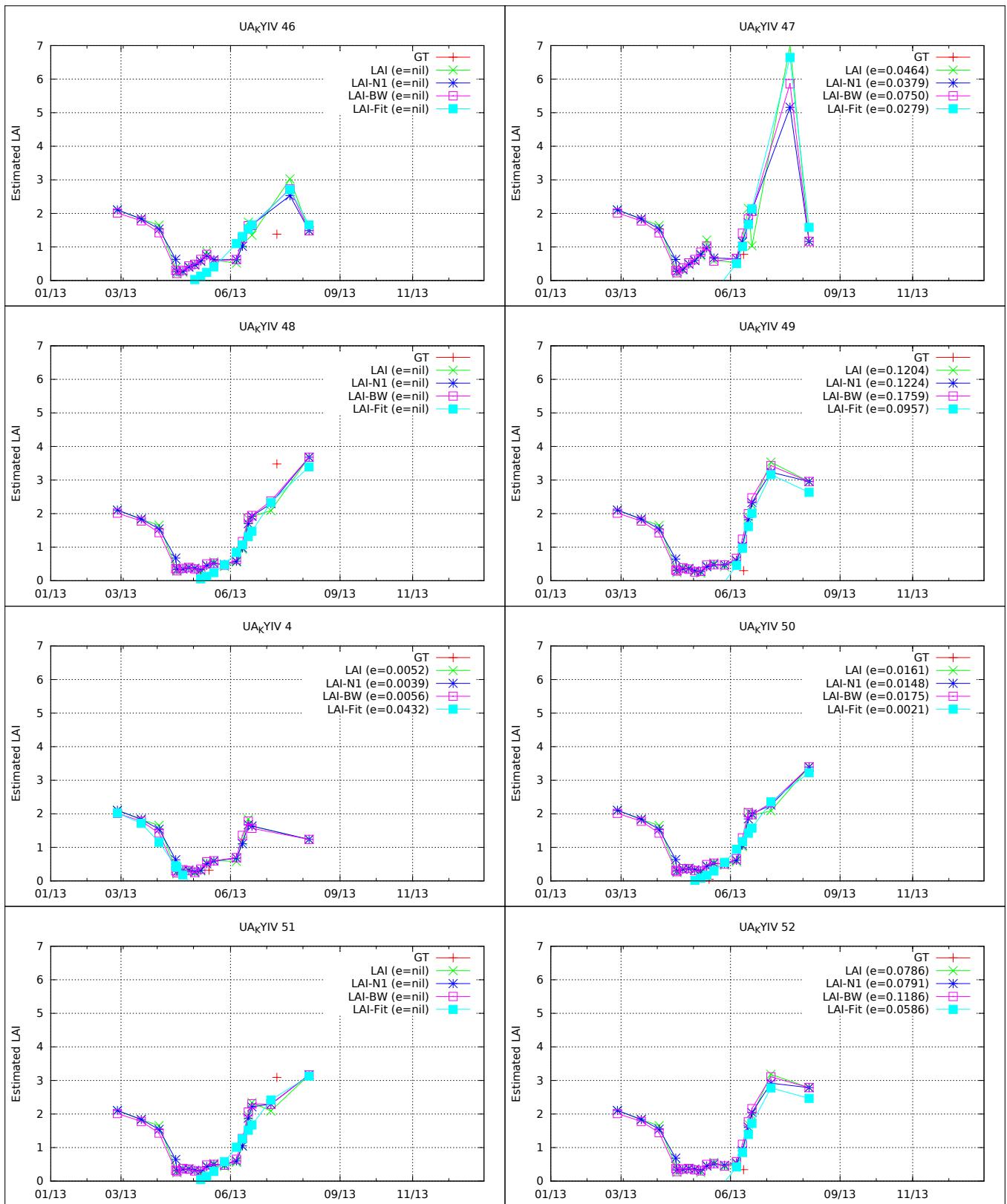


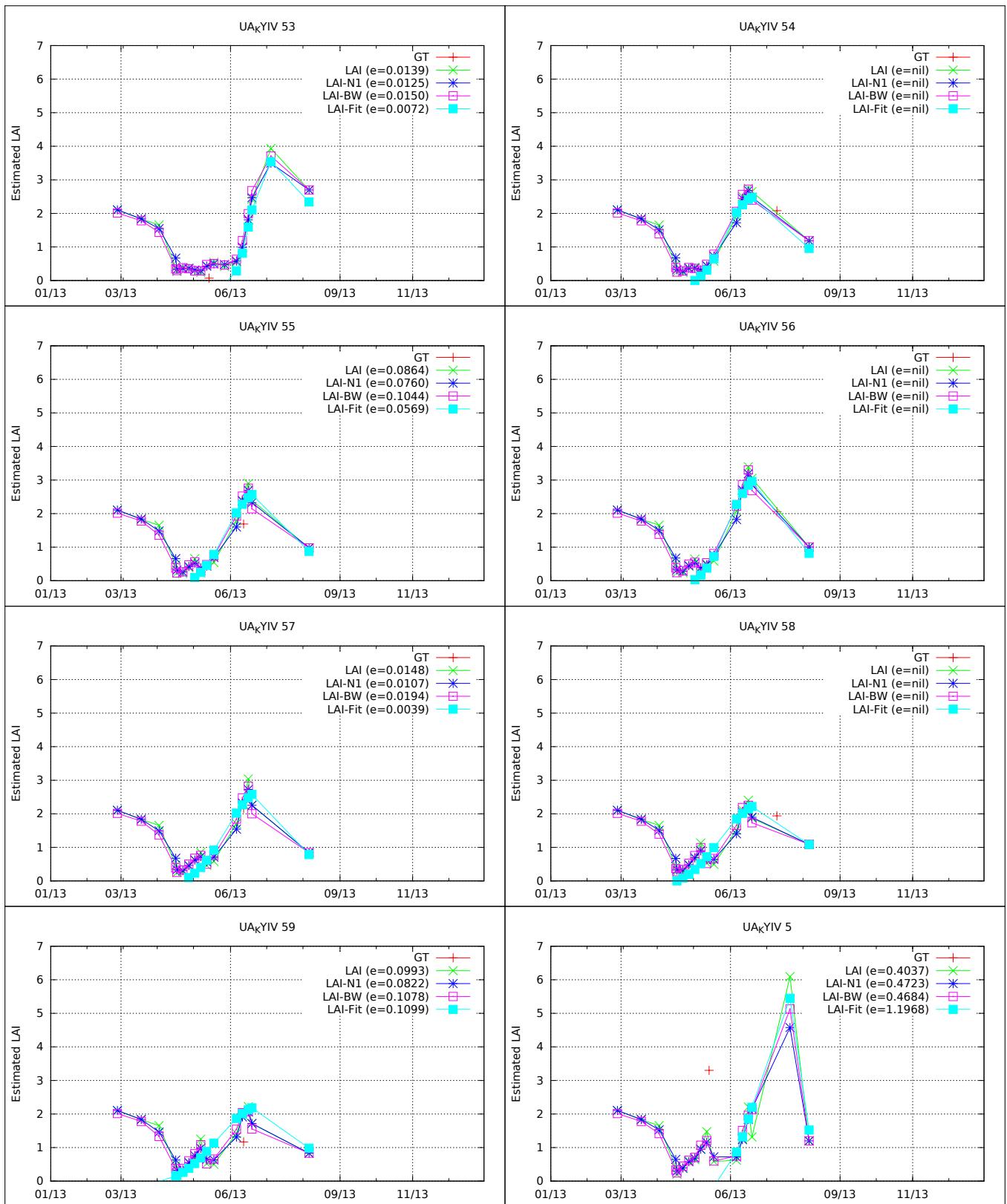


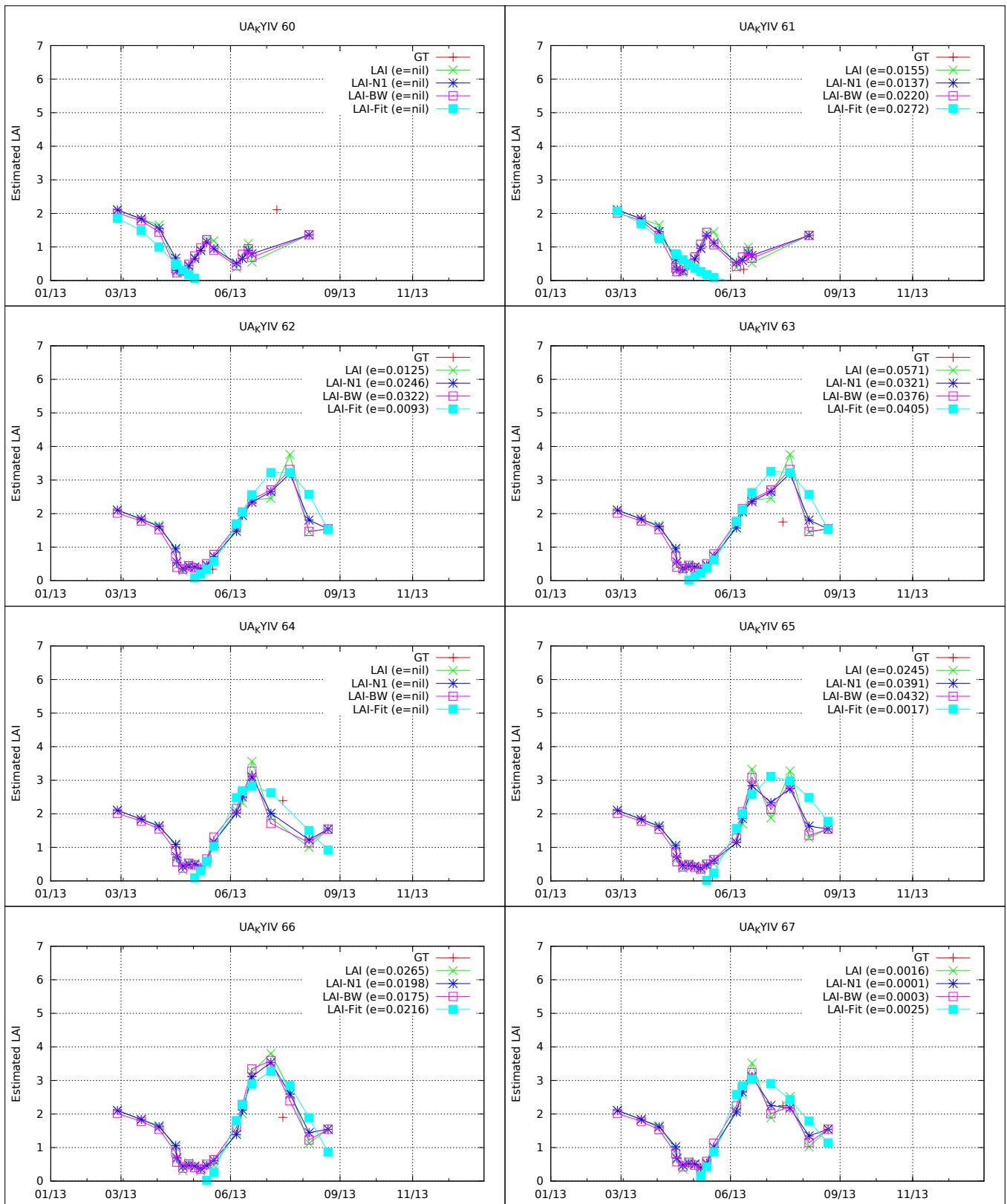


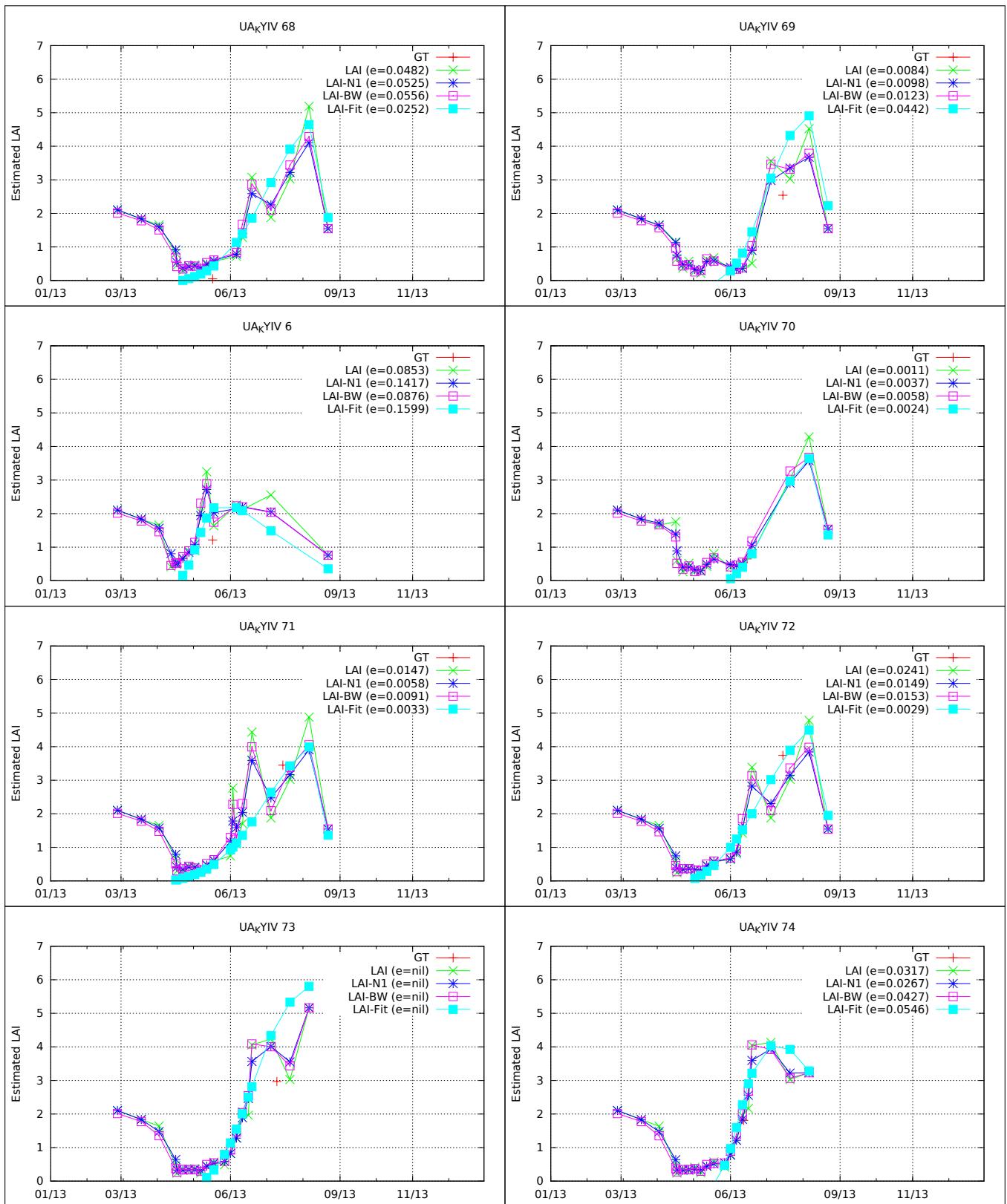


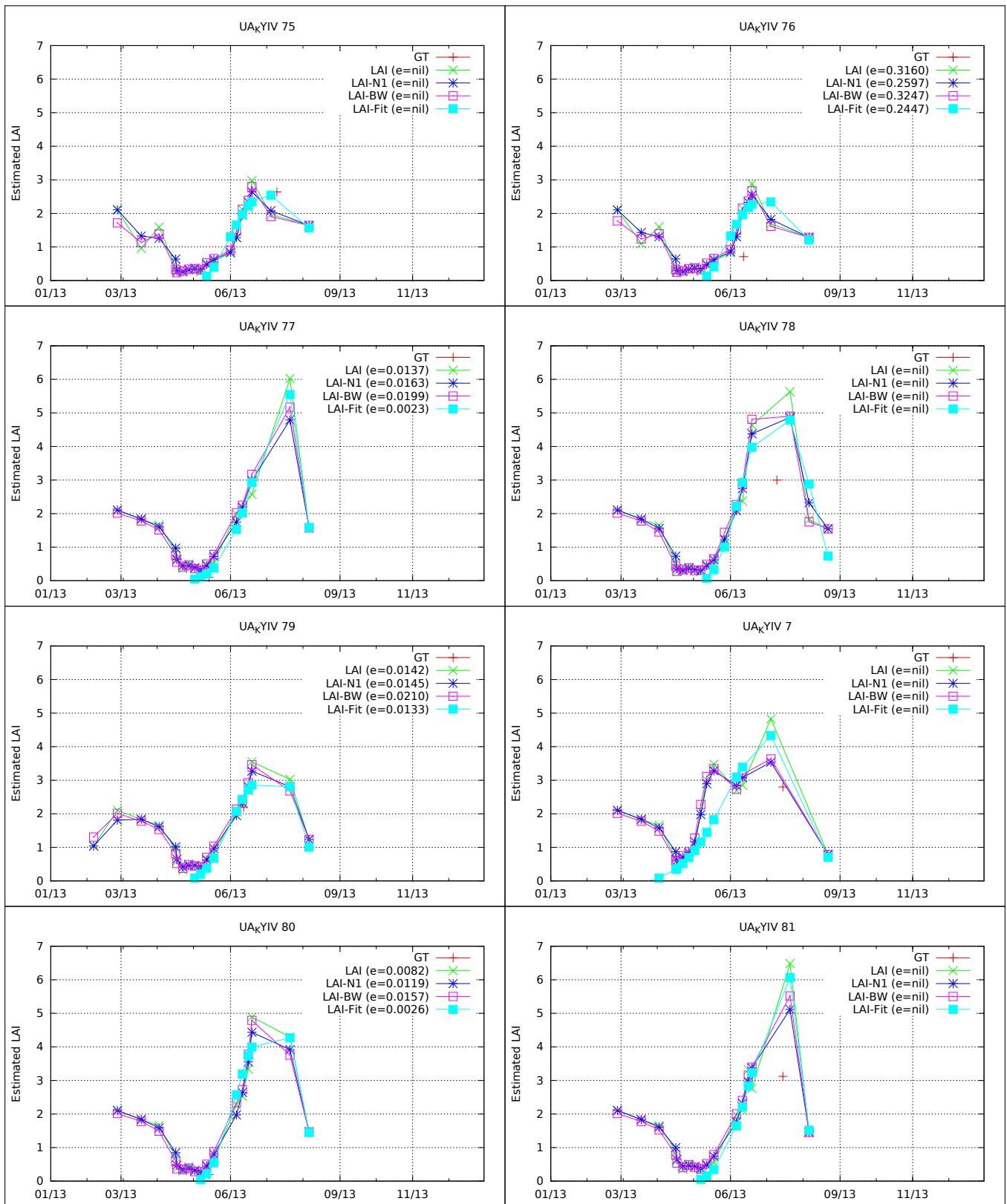


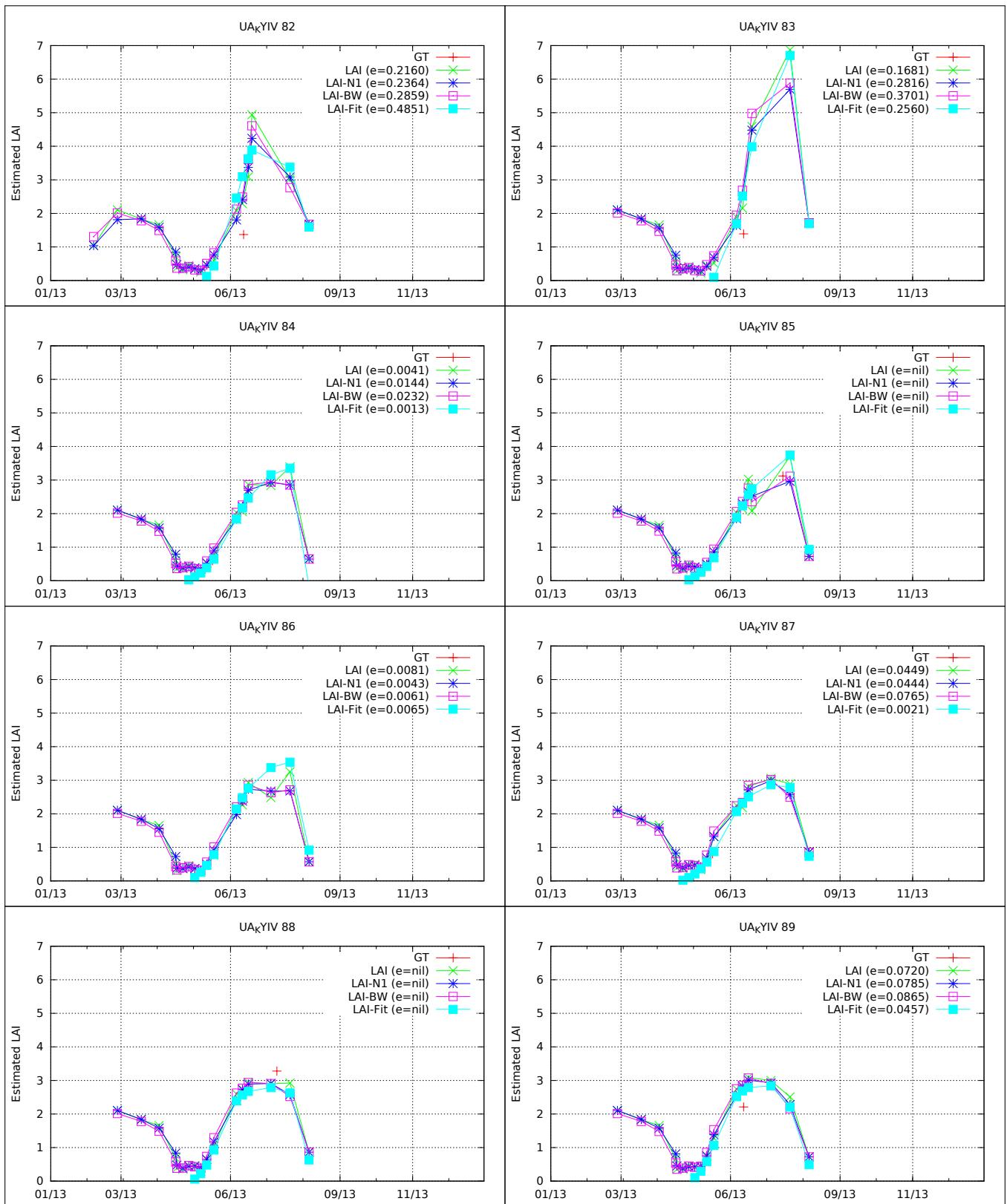


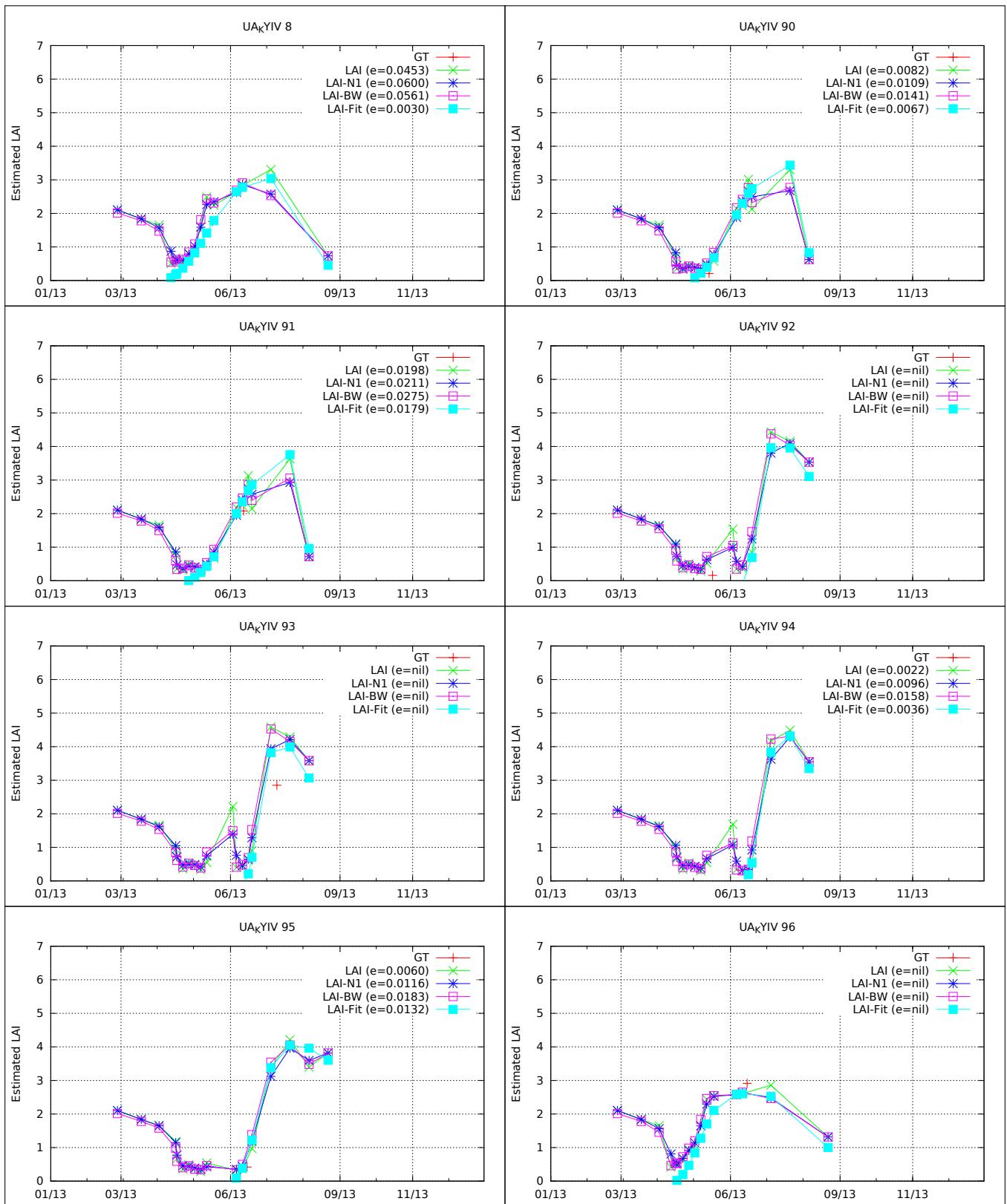


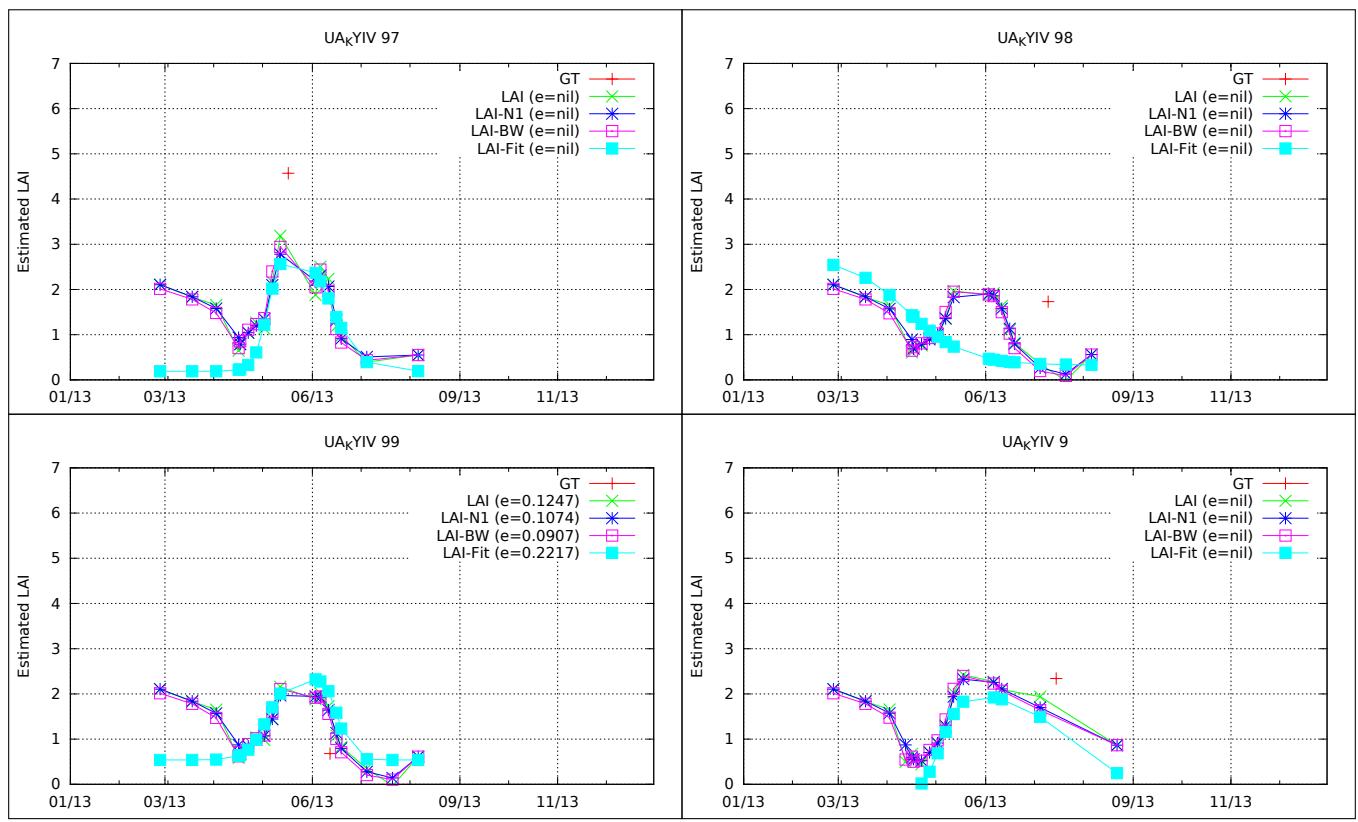












## 7.2 LAI images

The following pages illustrate the LAI estimation using the NLIRT M algorithm on a 20 km × 20 km area of the cloud-free images of the France - Sudmipy test site. Although the LAI estimation is performed for every pixel of the images, the result is combined with the cropland mask for these illustrations for an easier interpretation.

