



<b>Publication Year</b>	2021
<b>Acceptance in OA</b>	2023-05-18T10:51:14Z
<b>Title</b>	Solar surges related to UV bursts: Characterization through k-means, inversions and density diagnostics
<b>Authors</b>	Nóbrega Siverio, D., GUGLIELMINO, SALVATORE LUIGI, Sainz Dalda, A.
<b>Publisher's version (DOI)</b>	10.1051/0004-6361/202141472
<b>Handle</b>	<a href="http://hdl.handle.net/20.500.12386/34156">http://hdl.handle.net/20.500.12386/34156</a>
<b>Journal</b>	ASTRONOMY & ASTROPHYSICS
<b>Volume</b>	655

# Solar surges related to UV bursts

## Characterization through *k*-means, inversions, and density diagnostics<sup>★</sup>

D. Nóbrega-Siverio<sup>1,2,3,4</sup> , S. L. Guglielmino<sup>5,6</sup> , and A. Sainz Dalda<sup>7,8,9</sup> 

<sup>1</sup> Instituto de Astrofísica de Canarias, 38205 La Laguna, Tenerife, Spain  
e-mail: [dnobrega@iac.es](mailto:dnobrega@iac.es)

<sup>2</sup> Universidad de La Laguna, Dept. Astrofísica, 38206 La Laguna, Tenerife, Spain

<sup>3</sup> Rosseland Centre for Solar Physics, University of Oslo, PO Box 1029 Blindern, 0315 Oslo, Norway

<sup>4</sup> Institute of Theoretical Astrophysics, University of Oslo, PO Box 1029 Blindern, 0315 Oslo, Norway

<sup>5</sup> Dipartimento di Fisica e Astronomia “Ettore Majorana” – Sezione Astrofisica, Università degli Studi di Catania, Via S. Sofia 78, 95123 Catania, Italy

<sup>6</sup> INAF – Osservatorio Astrofisico di Catania, Via S. Sofia 78, 95123 Catania, Italy

<sup>7</sup> Lockheed Martin Solar & Astrophysics Laboratory, Palo Alto, CA 94304, USA

<sup>8</sup> Bay Area Environmental Research Institute, NASA Research Park, Moffett Field, CA 94035, USA

<sup>9</sup> Stanford University, HEPL, 466 Via Ortega, Stanford, CA 94305-4085, USA

Received 4 June 2021 / Accepted 31 August 2021

### ABSTRACT

**Context.** Surges are cool and dense ejections typically observed in chromospheric lines and closely related to other solar phenomena such as UV bursts or coronal jets. Even though surges have been observed for decades now, questions regarding their fundamental physical properties such as temperature and density, as well as their impact on upper layers of the solar atmosphere remain open.

**Aims.** Our aim is to address the current lack of inverted models and diagnostics of surges, as well as to characterize the chromospheric and transition region plasma of these phenomena.

**Methods.** We have analyzed an episode of recurrent surges related to UV bursts observed with the Interface Region Imaging Spectrograph (IRIS) in April 2016. The mid- and low-chromosphere of the surges were unprecedentedly examined by getting their representative Mg II *h&k* line profiles through the *k*-means algorithm and performing inversions on them using the state-of-the-art STIC code. We have studied the far-UV spectra focusing on the O IV 1399.8 Å and 1401.2 Å lines, which were previously unexplored for surges, carrying out density diagnostics to determine the transition region properties of these ejections. We have also used numerical experiments performed with the Bifrost code for comparisons.

**Results.** Thanks to the *k*-means clustering, we reduced the number of Mg II *h&k* profiles to invert by a factor 43.2. The inversions of the representative profiles show that the mid- and low-chromosphere of the surges are characterized, with a high degree of reliability, by temperatures mainly around  $T = 6$  kK at  $-6.0 \leq \log_{10}(\tau) \leq -3.2$ . For the electronic number density,  $n_e$ , and line-of-sight velocity,  $V_{\text{LOS}}$ , the most reliable results from the inversions are within  $-6.0 \leq \log_{10}(\tau) \leq -4.8$ , with  $n_e$  ranging from  $\sim 1.6 \times 10^{11} \text{ cm}^{-3}$  up to  $10^{12} \text{ cm}^{-3}$ , and  $V_{\text{LOS}}$  of a few  $\text{km s}^{-1}$ . We find, for the first time, observational evidence of enhanced O IV emission within the surges, indicating that these phenomena have a considerable impact on the transition region even in the weakest far-UV lines. The O IV emitting layers of the surges have an electron number density ranging from  $2.5 \times 10^{10} \text{ cm}^{-3}$  to  $10^{12} \text{ cm}^{-3}$ . The numerical simulations provide theoretical support in terms of the topology and location of the O IV emission within the surges.

**Key words.** Sun: atmosphere – Sun: chromosphere – Sun: transition region – methods: observational – methods: statistical

## 1. Introduction

Within the catalog of solar eruptive phenomena, surges are key chromospheric plasma ejections due to their frequent appearance as well as their close relationship with other important scientific targets of the Sun. The most common examples are flux emergence regions (e.g., Brooks et al. 2007; Guglielmino et al. 2010; Yang et al. 2018; Guglielmino et al. 2018, 2019a; Kontogiannis et al. 2020; Verma et al. 2020), Ellerman bombs (e.g., Watanabe et al. 2011; Yang et al. 2013; Vissers et al. 2013; Rutten et al. 2013), UV bursts (e.g., Kim et al. 2015; Huang et al. 2017; Rouppe van der Voort et al. 2017; Ortiz et al. 2020), coronal jets (e.g., Yokoyama & Shibata

1996; Canfield et al. 1996; Zhang & Ji 2014; Nelson et al. 2019; Ruan et al. 2019; Joshi et al. 2020a, 2021), and flares (e.g., McMath & Mohler 1948; Tandberg-Hanssen 1959; Roy 1973; Wang & Liu 2012; Huang et al. 2014; Schrijver & Higgins 2015), although there is also evidence of surges or surge-like ejections associated with light bridges (Asai et al. 2001; Shimizu et al. 2009; Robustini et al. 2016; Tian et al. 2018), explosive events (Madjarska et al. 2009), and eruptive filaments (Sterling et al. 2016; Li et al. 2017).

Mainly observed in H $\alpha$ , and in other lines such as Ca II 8542 Å, He I 10830 Å, H $\beta$  4861 Å, and Ca II H & K (Zhang et al. 2000; Liu & Kurokawa 2004; Nishizuka et al. 2008; Liu et al. 2009; Vargas Domínguez et al. 2014; Yang et al. 2014), it is now established, from both observations and numerical experiments, that surges are ejections typically with characteristic lengths of

<sup>★</sup> Movie associated with Fig. 2 is available at <https://www.aanda.org>

10–50 Mm (sometimes even reaching 200 Mm), they consist of small-scale thread-like features (Nelson & Doyle 2013; Li et al. 2016), they can be seen as a dense wall-like structure that appears to surround the emerged region (Moreno-Insertis et al. 2008; Moreno-Insertis & Galsgaard 2013), and their ejection direction depends on the particular geometry of the ambient magnetic field (MacTaggart et al. 2015). In addition, they can be recurrent (Jiang et al. 2007; Uddin et al. 2012; Wang et al. 2014) and/or have oscillations, as well as rotational and helical motions (Jibben & Canfield 2004; Jiang et al. 2007; Bong et al. 2014; Cho et al. 2019).

Despite this observational and theoretical effort, important questions concerning surges remain open. For instance, observational studies have mainly focused on their kinematic properties (see, e.g., Roy 1973; Gu et al. 1994; Chae et al. 1999; Chen et al. 2008; Ortiz et al. 2020), so fundamental quantities such as density and temperature, and their distribution with height, are not known. Recent simulations by Nóbrega-Siverio et al. (2016) using the Bifrost code (Gudiksen et al. 2011) point out that surges are composed of multi-thermal plasma populations that result from efficient heating and cooling mechanisms during their ejection. To provide observational support to these findings, it is necessary to properly derive constraints from chromospheric lines whose formation occurs in nonlocal thermodynamic equilibrium (NLTE) conditions and, moreover, which depend on radiative transfer effects such as partial frequency redistribution and 3D scattering (see Leenaarts et al. 2012); this is a complex task that is still pending. Another key aspect that remains poorly known is the impact of the surges on the transition region and corona. There are only a few examples in the literature where the relationship between transition region lines and surges has been studied, suggesting that a thin hot layer may surround the cooler  $H\alpha$  surge (see, e.g., Kirshner & Noyes 1971; Schmieder et al. 1984; Madjarska et al. 2009). Combining theory and observations, Nóbrega-Siverio et al. (2017, 2018) show that surges have enhanced emission on Si IV which is related to nonequilibrium ionization effects during the surge formation; nonetheless, the Interface Region Imaging Spectrograph (IRIS; De Pontieu et al. 2014) observations analyzed by these authors only have spectral information at the base of the surge due to the location of the raster. Independent observational evidence of enhanced Si IV emission associated with surges with IRIS was found by Guglielmino et al. (2019a). Nevertheless, none of the aforementioned authors have performed diagnostics to extract valuable information on the hottest regions of the surges and the weakest lines of the transition region are still unexplored.

The purpose of this work is to tackle the present lack of inverted models and diagnostics of surges, filling the gap in the knowledge of important physical quantities of these key phenomena for the solar atmosphere. To that end, we analyze a series of surges observed by IRIS within Active Region (AR) NOAA 12529 (Guglielmino et al. 2017, 2019b). For the first time, the characterization of the chromospheric and transition region properties of the surge plasma has been achieved by combining machine learning techniques, inversions, and density diagnostics. In addition, realistic simulations were used to provide theoretical support to the observations. The layout of this paper is as follows. In Sect. 2, we describe the observational data, the methods, and the numerical experiments employed. In Sect. 3, we study the surges focusing on the Mg II  $h&k$  line (Sect. 3.1), the O IV lines (Sect. 3.2), and the comparisons with the simulations (Sect. 3.3). Finally, Sect. 4 contains the discussion and main conclusions.

## 2. Observations, methods, and numerical experiments

### 2.1. Observational data

For this paper, we have used IRIS observations obtained in 2016 between 22:34:43 UT on April 13 and 01:55:29 UT on April 14, within the active region AR NOAA 12529, which passed on the solar disk exhibiting an almost typical bipolar  $\beta$  configuration at heliocentric coordinates  $(X, Y) \approx (0'', 260'')$  (central solar meridian passage) and heliocentric angle  $\mu = 0.96$ . In particular, we focused on the recurrent surges that appear as a consequence of magnetic flux emergence. The magnetic flux emergence episode, which originated in the diffuse plage of the following positive polarity of the AR, as well as the long-lasting (3 h) UV bursts were analyzed in detail by Guglielmino et al. (2018, 2019a).

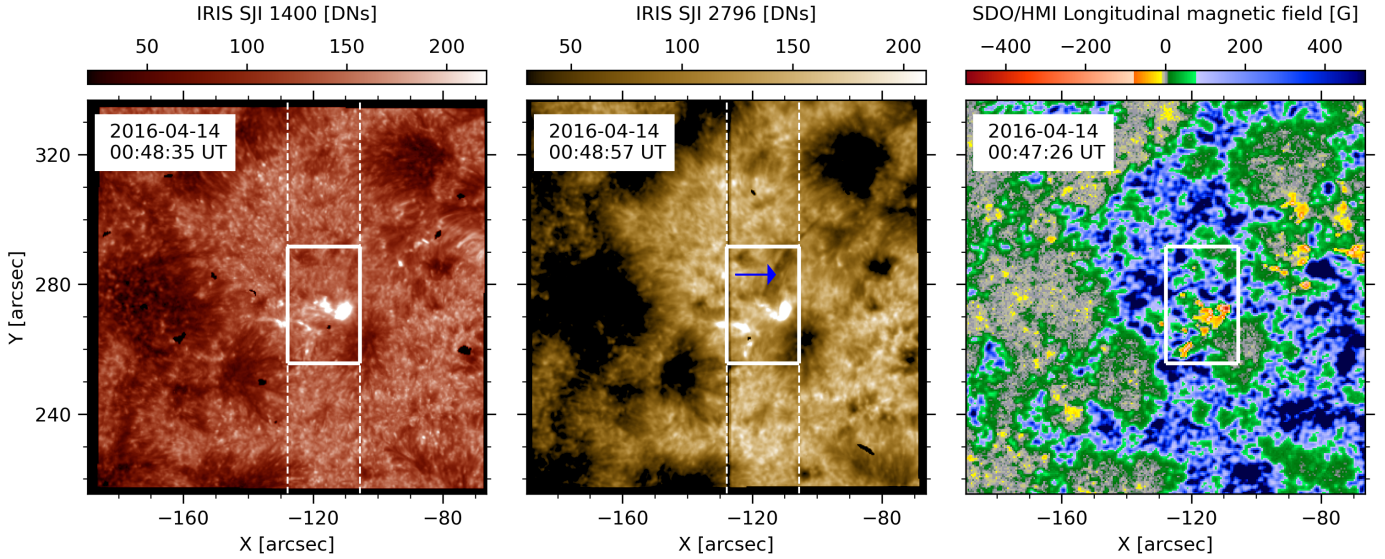
The IRIS data set corresponds to the program OBS3610113456 whose sequence consists of six dense 64-step raster scans. In this paper we have employed a more recent calibration (version 1.84) of the IRIS level 2 data products than the one originally applied by Guglielmino et al. (2018, 2019a) (version 1.56) due to updates carried out by the instrument team. In addition, we only used the last four scans of the sequence due to a change in the exposure time during the observations, maintaining the notation previously used to identify them, namely, Raster 3, 4, 5, and 6. The selected raster scans contain UV spectra acquired in seven spectral ranges, encompassing C II 1334.5 and 1335.7 Å, Si IV 1394 and 1402 Å, Mg II  $k$  2796.3 and  $h$  2803.5 Å lines, and the faint lines around the chromospheric O I 1355.6 Å line. The exposure time is 18 s for the far-UV (FUV) channels and 9 s for the near-UV (NUV) channel. The sequence has a pixel size of  $0''.35$  along the  $y$  direction, with a  $0''.33$  step size along the  $x$  direction, and the covered field of view (FoV) was  $22''.3 \times 128''.4$ . Slit-jaw images (SJIs) were simultaneously acquired in the 1400 and 2796 Å passbands, with a cadence of 63 s. In addition, to link the IRIS observations to the photospheric magnetic field, we employed data from the Helioseismic and Magnetic Imager (HMI; Scherrer et al. 2012) on board the Solar Dynamics Observatory (SDO; Pesnell et al. 2012). For further details about the IRIS/SDO data set, we refer the reader to Guglielmino et al. (2018, 2019a).

Figure 1 displays the context of the observations at  $\sim 00:48$  UT on 2016 April 14. The left panel contains the IRIS SJI 1400 Å, showing a strong brightening (a UV burst) located around  $X = -110''$  and  $Y = 272''$ . In the middle panel, which illustrates the IRIS SJI 2796 Å, the UV burst is visible next to an elongated dark structure that corresponds to a surge (blue arrow). The right panel contains the SDO/HMI line-of-sight (LOS) magnetogram, showing the diffuse plage that follows the positive polarity of the AR. At the time of the image, a flux emergence episode was taking place and a negative polarity patch was clearly distinguishable embedded in the positive polarity dominated region. In the figure, the portion of the FoV analyzed in this paper is delimited by a solid box encompassing a sub-FoV of  $22''.3 \times 35''.9$ , which ranged from  $X = -127''.9$  to  $X = -105''.6$  and  $Y = 255''.7$  to  $Y = 291''.6$  at the time of the image.

### 2.2. Methods

#### 2.2.1. The $k$ -means algorithm

The  $k$ -means clustering is an unsupervised algorithm, popular in data-mining, machine learning, and artificial intelligence, which



**Fig. 1.** Context image of the events associated with the flux emergence region in the plage following the positive polarity of the AR NOAA 12529 at ~00:48 UT on 2016 April 14. *Left:* IRIS SJI 1400 Å. *Middle:* IRIS SJI 2796 Å. *Right:* SDO/HMI line-of-sight (LOS) magnetogram. The vertical dashed lines indicate the area covered by the IRIS slit during the raster scans. The solid box, ranging from  $X = -127''9$  to  $X = -105''6$  and  $Y = 255''7$  to  $Y = 291''6$ , frames the portion of the FoV studied in this paper. The blue arrow indicates one of the observed surges. The axes give the distance from the solar disk center.

classifies a set of  $n$  samples in  $k$  disjoint groups (or clusters) of equal variance. To that end, it minimizes the within-cluster sum-of-squares (also known as inertia) given by

$$\sum_{i=1}^n \sum_{j=1}^k \min \left( \|x_i - \mu_j\|^2 \right), \quad (1)$$

where  $x_i$  is the  $i$ -th observed point and  $\mu_j$  is the mean of the  $j$ -th cluster center or centroid.

This algorithm is relatively easy to implement, scales to large data sets, and can be adapted to group line profiles. In solar physics,  $k$ -means has gained popularity being used, for instance, to classify Stokes profiles and chromospheric lines in different phenomena such as flares and spicules (see, e.g., Pietarila et al. 2007; Viticchié & Sánchez Almeida 2011; Panos et al. 2018; Sainz Dalda et al. 2019; Bose et al. 2019; Robustini et al. 2019; Joshi & Rouppe van der Voort 2020b; Kuckein et al. 2020; Bose et al. 2021a,b; Barczynski et al. 2021).

In our case, we used it to classify the Mg II  $h&k$  profiles obtained by IRIS within the sub-FoV shown as a solid box in Fig. 1. This way, the  $k$ -means method can help us to identify whether surges have characteristic Mg II  $h&k$  line profiles that may indicate similar atmospheric properties. The reason to classify the spectra is inspired by the theoretical work by Nóbrega-Siverio et al. (2016), in which the authors discern different populations within the surge depending on the thermodynamics and evolution of the plasma. Details about the data processing used to feed the  $k$ -means algorithm as well as the choice of the number of clusters  $k$  can be found in Appendix A.

### 2.2.2. Inversions

Inversions of the Mg II  $h&k$  lines have been carried out using the STockholm inversion Code<sup>1</sup> (STiC, de la Cruz Rodríguez et al. 2019, 2016), which assumes NLTE and includes partial frequency redistribution effects of scattered elec-

trons (Leenaarts et al. 2012). This code allows us to get a stratified model of the solar atmosphere that covers, depending on the spectral lines considered to be inverted, the photosphere, the chromosphere, and the transition region. An initial guess model atmosphere is slightly modified in an iterative process until the output radiation – obtained by solving the radiative transfer equation taking the considerations mentioned above into account – fits the observed spectra. Readers interested in the problem of the inversion of spectral data in the context of solar physics can find an excellent review by del Toro Iniesta & Ruiz Cobo (2016).

We also explored the possibility of inverting the Mg II  $h&k$  surge profiles using IRIS<sup>2</sup> (Sainz Dalda et al. 2019)<sup>2</sup>. Its inversion scheme (i.e., the number of nodes used in the thermodynamics variables to vary the model atmosphere at different optical depths) consists of the following two cycles: in the first cycle, it uses four nodes in temperature, and three nodes both in the LOS and turbulent velocities; in the second cycle, it uses seven nodes in temperature and four nodes in both velocities. Even though this approach is good enough for a large variety of profiles, it struggles with profiles with extreme line widths, such as those observed in X-class flares and UV bursts, as well as with profiles related to eruptive events. This means that the current version of IRIS<sup>2</sup> is not appropriated for phenomena such as surges.

Taking the above mentioned point into consideration, we proceeded to invert the Mg II  $h&k$  profiles of our observations with a more complex scheme than the one used to build the IRIS<sup>2</sup> database. After different tests, we selected an inversion scheme that consists of four cycles whose details are summarized in Table 1. For cycles 1 and 2, the number of nodes are the same as those used to build the IRIS<sup>2</sup> database. The temperature,  $T$ , and electron density,  $n_e$ , in the initial guess model atmosphere for the cycle 1 are taken from the FALC model (Fontenla et al. 1993). The micro-turbulence velocity,  $V_{\text{turb}}$ , and the LOS velocity,  $V_{\text{LOS}}$ , were introduced ad hoc as a smooth gradient along the optical depth between  $-6 \leq \log_{10}(\tau) \leq 0$  that goes from 3 to 1 km s<sup>-1</sup> for

<sup>1</sup> <https://github.com/jaimedelacruz/stic>

<sup>2</sup> <https://iris.lmsal.com/iris2/>

**Table 1.** Number of nodes per cycle considered for the inversion of the Mg II *h&k* spectral profiles in this article.

Nodes in variable for cycle #	1	2	3	4
$T$	4	7	9	9
$V_{\text{turb}}$	3	4	7	9
$V_{\text{LOS}}$	3	4	4	6

$V_{\text{turb}}$ , and from 1 to  $0.5 \text{ km s}^{-1}$  for  $V_{\text{LOS}}$ . The initial model atmosphere for the cycles 2, 3, and 4 is the output model atmosphere obtained from the previous cycle. The computation of the inversion uncertainties,  $\sigma$ , was carried out following Eq. (42) from the paper by [del Toro Iniesta & Ruiz Cobo \(2016\)](#); the same calculation was performed by [Sainz Dalda et al. \(2019\)](#) for the IRIS<sup>2</sup> database.

### 2.2.3. Density diagnostics

The density diagnostics method is based on the theoretical relationship between the electron density and intensity ratio of the O IV 1399.8 Å and 1401.2 Å lines. Indeed, these intercombination lines are known to provide useful electron density diagnostics in a variety of solar features and astrophysical plasmas (see, e.g., [Feldman & Doschek 1979](#)) because their ratios are largely independent of the electron temperature, and only weakly dependent on the electron distribution ([Dudík et al. 2014](#)). A recent discussion about the diagnostic potential of these lines in IRIS observations has been presented by [Polito et al. \(2016a\)](#), who also show that their sensitivity interval for density diagnostics range between  $10^{10} - 10^{12} \text{ cm}^{-3}$ .

The theoretical intensity ratios between the O IV 1399.8 Å and 1401.2 Å lines were used to derive the electron density from the observed line ratio, under the assumption of ionization equilibrium (CHIANTI v.8 database, [Del Zanna et al. 2015](#)), with a formation temperature of  $\log_{10}(T) = 5.15$  K. We obtained the line peak intensity from each considered pixel by fitting a Gaussian function to the O IV 1399.8 Å and 1401.2 Å line profiles, respectively, also retrieving the uncertainties of the intensities. Blends in these emission lines are negligible in our study, as they mostly occur during flares ([Polito et al. 2016b](#)). Finally, for the computation of the density, we used the IDL *SolarSoft* routine `iris_ne_oiv.pro`.

### 2.3. Numerical experiments

To get a joint perspective between observations and theory, we used three different 2.5D numerical experiments in which surges were ejected as a result of magnetic flux emergence. The simulations were performed with the radiation-magnetohydrodynamic Bifrost code ([Gudiksen et al. 2011](#)) assuming LTE and including thermal conduction along the magnetic field lines, optically thin losses, and a radiation transfer adequate for the photosphere and chromosphere. The details of these simulations have been described in the papers by [Nóbrega-Siverio et al. \(2016, 2017, 2018\)](#).

## 3. Results

For the purpose of showing the surges analyzed in this paper, the left column of Fig. 2 contains radiance maps for the chromospheric Mg II k 2796.3 Å line within the selected sub-FoV (solid

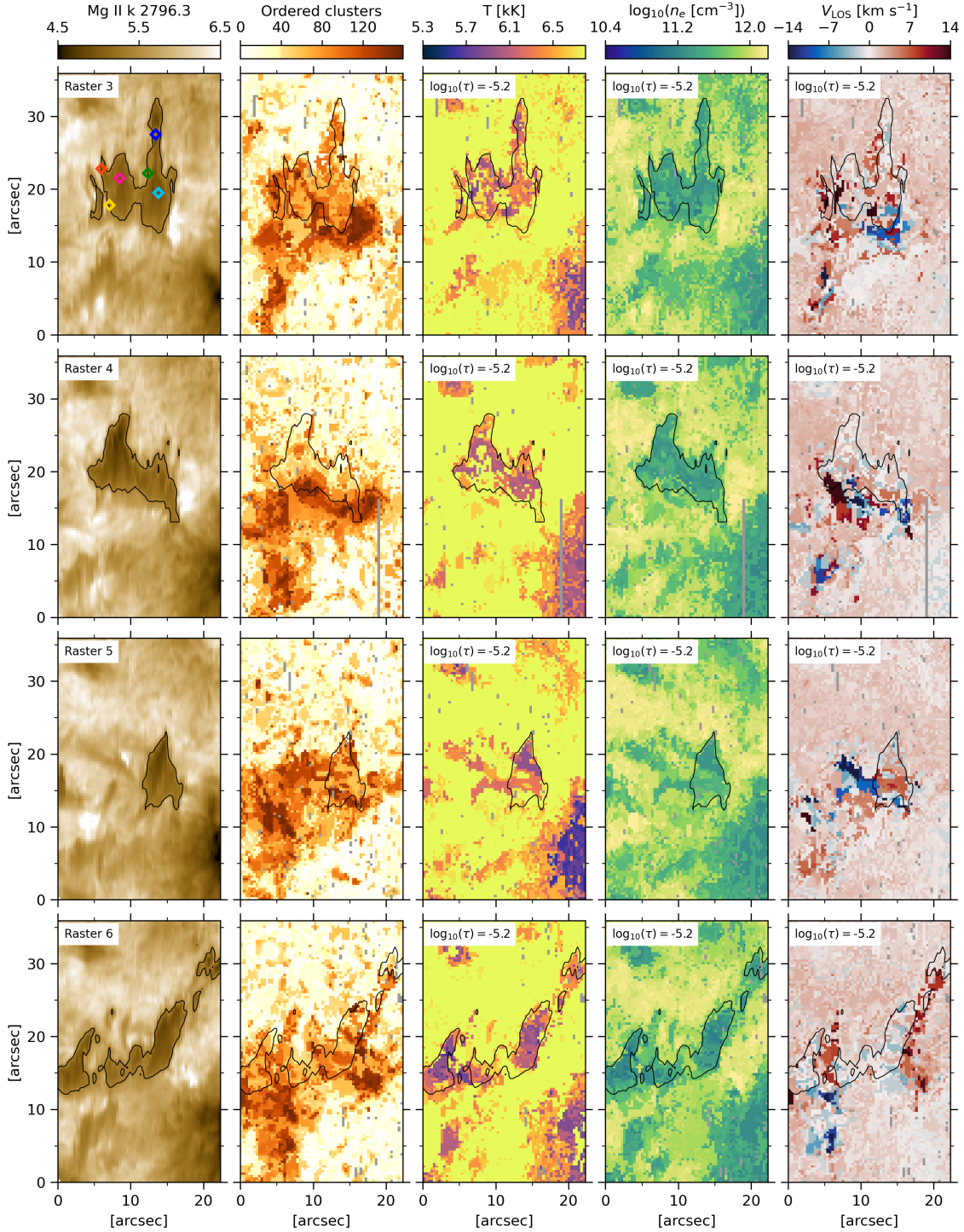
box in Fig. 1) during the four IRIS raster scans. These maps were computed by averaging the intensity within  $\pm 15 \text{ km s}^{-1}$  with respect to the core of the aforementioned line. In the panels, we added contours to ease the location of the bulk of the surges, which can be recognized by their dark and elongated structures composed of threads. The surges can be more extended spatially than the regions limited by these contours since they only refer to the radiance maps in the center of Mg II k 2796.3 Å. In the panels, it is also possible to identify the location of the UV bursts associated with the surges, especially in the third and fourth rasters, due to their enhanced brightness (see [Guglielmino et al. 2019a](#), for details about these UV bursts). In the following, we focus on the observed surges by studying their chromospheric properties through the Mg II *h&k* line (Sect. 3.1), analyzing their transition region counterpart using the O IV 1399.8 Å and 1401.2 Å lines (Sect. 3.2), and comparing the observational results with the simulations (Sect. 3.3).

### 3.1. Analysis of Mg II *h&k* in the surges

#### 3.1.1. Mg II *h&k* clustering and profile examples

The first step prior to the analysis of the Mg II *h&k* spectra within the surges is the clustering through the *k*-means method (Sect. 2.2.1). In addition to identifying representative profiles in the phenomenon of interest, this method allowed us to noticeably reduce the number of Mg II *h&k* profiles that would be necessary to analyze in principle. In our particular case, we chose  $k = 160$  for each raster, thus getting a total of 640 representative profiles that were later inverted. This implies a reduction by a factor 43.2 with respect to the original total number of profiles (27 648) within the IRIS sub-FoV chosen for this paper. The second column of Fig. 2 contains the label maps of the clusters for each raster. The labels were assigned to the clusters depending on the number of profiles that compose each clusters: the lower the label number, the greater the number of profiles in the corresponding cluster. By ordering the labels, we can already discern that surges and their surroundings have clearly different Mg II *h&k* profiles from other regions.

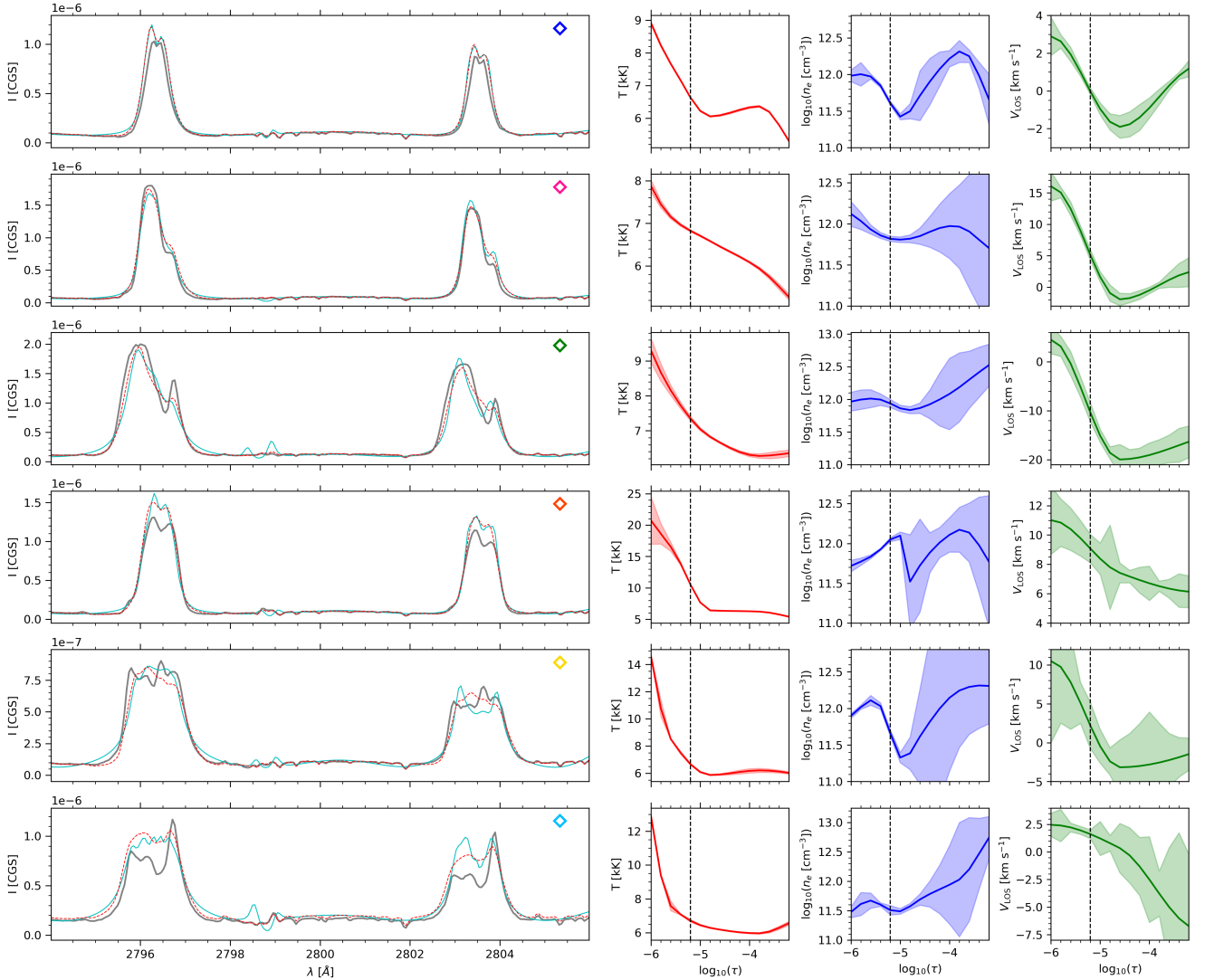
In order to show the peculiarity of some of the surge profiles and to compare them with the representative ones obtained with the *k*-means, we selected six different pixels in the surge of Raster 3 (colored diamonds in Fig. 2), which are located in the following places: an elongated finger-like thread (blue); the left side of the bulk of the surge (pink); some particular regions at the boundaries of the ejection (orange, green, and yellow); and the right side of the bulk of the surge (cyan). The last four locations are also used in the analysis of the O IV lines (Sect. 3.2). The Mg II *h&k* spectral profiles for these six locations are plotted in the left column of Fig. 3 as gray lines. They show a large diversity, for instance, most of the profiles have greater  $k_{2V}$  and  $h_{2V}$  peaks than  $k_{2R}$  and  $h_{2R}$ , respectively, except in the cyan location, which suggests different Doppler motions in the upper chromosphere ([Leenaarts et al. 2013](#)). It is also possible to find  $k_3$  and  $h_3$  ranging from clearly visible valleys (as in the green location) to an almost flat behavior (blue pixel); broad profiles with several peaks (yellow and cyan); and an interesting absence of emission in the Mg II triplet in all of them, which is contrarily observed in UV bursts ([Hong et al. 2017](#)). In the panels, we also plotted, as a red dashed line, the corresponding representative profiles obtained with the *k*-means method. The representative profiles match the observational data quite well, particularly for the blue, pink, orange, and green locations, which only show slight differences in the intensity while the shape of the Mg II *h&k* lines



**Fig. 2.** Observed surges and results from  $k$ -means and inversions. *Columns from left to right:* radiance maps in the core of the Mg II k 2796.3 Å line; cluster labels from the  $k$ -means ordered by the amount of profiles within a cluster; and maps at  $\log_{10}(\tau) = -5.2$  for temperature,  $T$ , electron number density,  $n_e$ , and LOS velocity,  $V_{\text{LOS}}$ , from the inversions of the Mg II  $h\&k$  line using the STiC code. Black contours delimit the bulk of the surges visible in the Mg II k 2796.3 Å radiance maps. Bad pixels and pixels affected by cosmic rays are masked and shown with a color gray. Colored diamonds are superimposed in the first panel for later reference. An animation of this figure is available [online](#) varying  $\log_{10}(\tau)$  from  $-6.0$  to  $-3.2$ .

and the continuum are accurately represented. In the case of the cyan and yellow locations, the continuum and the width of Mg II  $h\&k$  lines are properly captured with the representative profile; nonetheless, the shape of the peaks of the lines is less similar. We need to highlight that this discrepancy is a natural result expected from  $k$ -means: there are always some observed profiles

that are assigned to the cluster with the most similar profiles, but still they are not nicely represented. In our case, we checked that these divergent profiles represent a minority within the clusters (we validated the good behavior of the  $k$ -means algorithm by three different methods as mentioned in the Appendix A). Consequently, this small number of outliers have little impact on our



**Fig. 3.** Spectral information and computed atmospheres for the six positions indicated with colored diamonds in Fig. 2. *Left column:* Observed Mg II  $h&k$  profile (gray line), the corresponding representative profile obtained with the  $k$ -means method (red dashed line), and the inverted profile using the STiC code (cyan line) in CGS units ( $\text{erg cm}^{-2} \text{s}^{-1} \text{sr}^{-1} \text{Hz}^{-1}$ ). *Three rightmost columns:* results from the inversions for  $T$  (red),  $n_e$  (blue), and  $V_{\text{LOS}}$  (green) as functions of  $\log_{10}(\tau)$  with their corresponding inversion uncertainties ( $\pm\sigma$ ). The vertical dashed line in these plots at  $\log_{10}(\tau) = -5.2$  indicates the optical depth shown in Fig. 2.

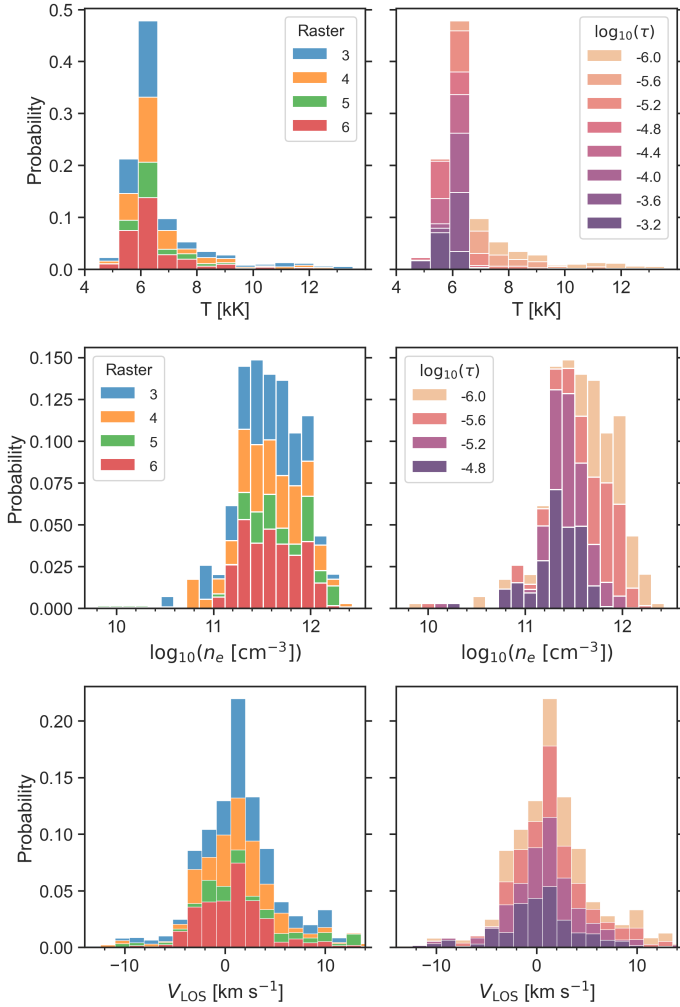
results since we are interested in an overall view of the physical properties derived from inversions as discussed in the following section.

### 3.1.2. Inversions of the representative Mg II $h&k$ profiles and statistics

Once we have the representative profiles for each raster, we proceeded with their corresponding inversions using the state-of-the-art STiC code with the scheme explained in Sect. 2.2.2. The third, fourth, and fifth columns of Fig. 2 contain the results for the temperature,  $T$ , the electronic number density,  $n_e$ , and the LOS velocity,  $V_{\text{LOS}}$ , respectively, at  $\log_{10}(\tau) = -5.2$ . In the maps, it is possible to perceive that the surges have some peculiarities in their physical properties, for example, they are mostly cooler and with a smaller electron number density than their surroundings. With respect to the velocity, the maps reveal that the bulk of the surges have values of a few  $\text{km s}^{-1}$ , while the largest values are located next to the edges of the contours

of the finger-like threads (see Raster 3 and 6), reaching up to  $\sim 11 \text{ km s}^{-1}$ . An animation of this figure is also available, showing the variation of the properties with the optical depth between  $-6.0 \leq \log_{10}(\tau) \leq -3.2$ .

In the panels of the left column of Fig. 3, we plotted the inverted profiles for the selected locations of Raster 3 as cyan lines to compare them with the representative profiles obtained with the  $k$ -means method. This way, we can get an idea about whether the inversion scheme used in this paper (Table 1) is adequate for such complex profiles. We have found that, in most cases, the inverted profiles mimic the representative profiles with a high grade of detail and that the major discrepancies are in the peculiar profiles of the cyan and yellow locations. The right panel of Fig. 3 contains the corresponding representative model atmospheres extracted from the selected locations to illustrate the variation of  $T$  (red),  $n_e$  (blue), and  $V_{\text{LOS}}$  (green) with  $\log_{10}(\tau)$ . For each physical quantity, we also superimposed the inversion uncertainties ( $\pm\sigma$ ). The uncertainties are small for the temperature in the whole range of optical depths, while



**Fig. 4.** Statistical results for  $T$  (top row),  $n_e$  (middle row), and  $V_{\text{LOS}}$  (bottom row) obtained within the contour that delimits the bulk of the different surges in Fig. 2. The statistics contain data from the optical depths where inversions are more reliable, from  $\log_{10}(\tau) = -6.0$  to  $-3.2$ , for the temperature, and from  $\log_{10}(\tau) = -6.0$  to  $-4.8$  for the density and velocity. *Left column:* stacked histograms organized by raster, showing that the different surges have similar properties. *Right panel:* stacked histograms organized by  $\log_{10}(\tau)$  illustrating the variation of the physical parameters with the optical depth.

for the density, the smaller uncertainties are located higher up in the atmosphere (more negative values of  $\log_{10}(\tau)$ ). The latter is also true for the velocity in the blue, pink, green, and cyan locations. We inspected the behavior of the uncertainties for all the results within the black contour that delimits the bulk of the different surges in Fig. 2, concluding the following: for the temperature, the accuracy of the inversions is excellent for all the optical depths within  $\log_{10}(\tau) = -6.0$  and  $-3.2$ ; for the density, the inversions' results can be rather acceptable between  $\log_{10}(\tau) = -6.0$  and  $-4.8$ ; and for the velocities, the range between  $\log_{10}(\tau) = -6.0$  and  $-4.8$  is also the best one generally.

To get a whole perspective of the general properties of the surges, we need to perform a statistical analysis. Figure 4 contains histograms for  $T$  (top row),  $n_e$  (middle row), and  $V_{\text{LOS}}$  (bottom row), obtained within the black contour that delimits the bulk of the different surges in Fig. 2. To create these histograms, we considered the range of optical depths in which

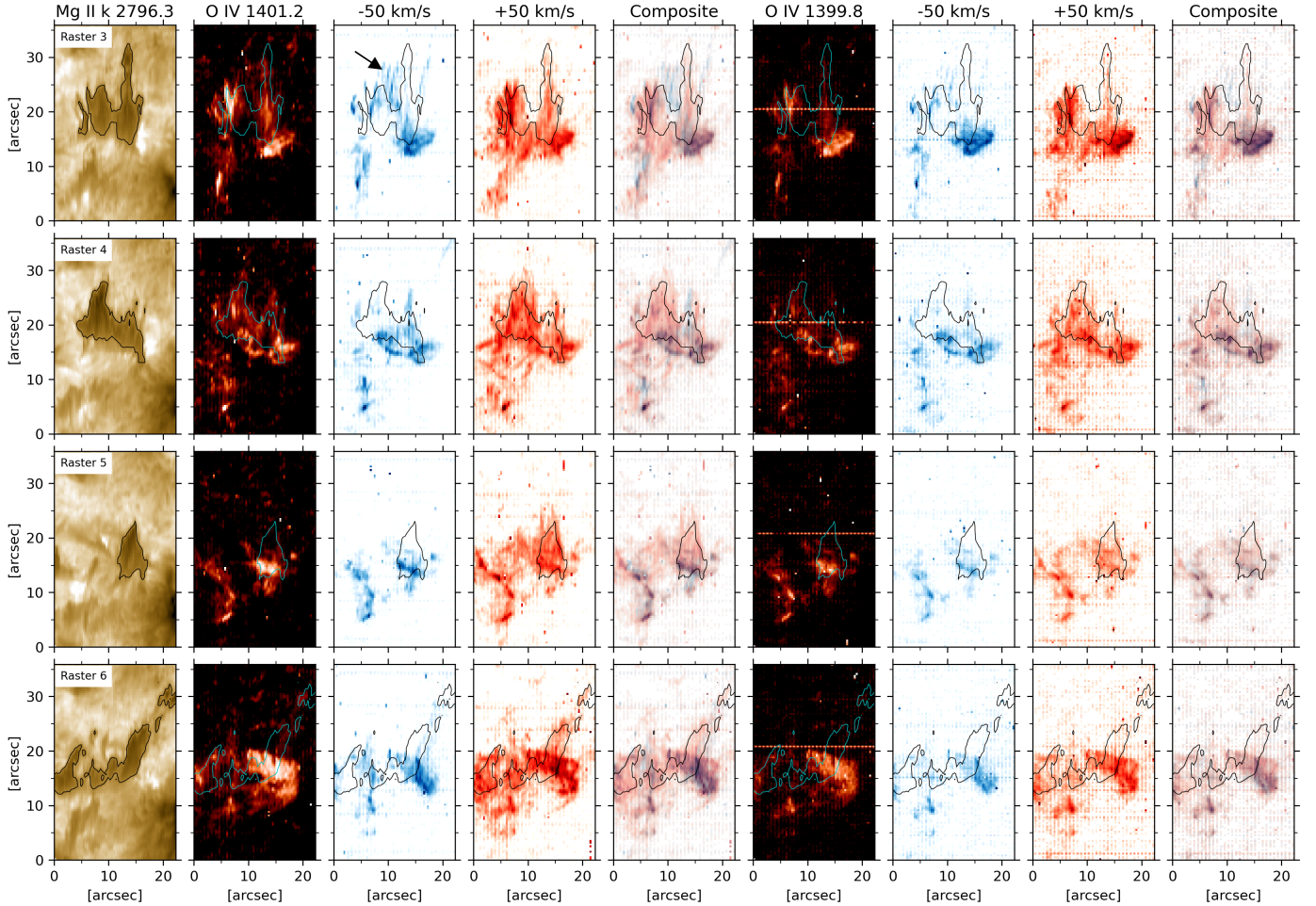
the uncertainties are smaller for each physical quantity. The histograms in the left column of the figure are stacked by rasters, that is, each bar in the chart represents the whole probability, and the length of the segments in the bar represents the contribution from each raster. Even though we performed  $k$ -means individually for each raster with topologically different surges, the statistical distributions for the temperature, density, and velocity that we obtained after the inversions of the representative profiles are very similar: an important result that confers robustness to our method to characterize the properties of the surges. The characteristic properties of our surges are as follows: the most probable temperature is around 6 kK, the electronic number density is mostly concentrated from  $\sim 1.6 \times 10^{11}$  to  $10^{12} \text{ cm}^{-3}$ , and the LOS velocity is typically of a few  $\text{km s}^{-1}$ . The right column of Fig. 4 shows the same statistics, but stacked by the optical depth, highlighting the variation of the parameters with the optical depth. In this case, cooler plasma with a smaller electron number density in the surges is located at less negative values of the optical depth, that is, deeper in the atmosphere.

### 3.2. Analysis of O IV in the surges

#### 3.2.1. Enhanced O IV emissivity within the surge

One of the most striking results is the first finding of emission in the O IV 1401.2 Å and O IV 1399.8 Å lines in surges. To illustrate this fact, Fig. 5 contains the Mg II  $k$  2796.3 Å line radiance maps for the four rasters (first column), to show the context, and the corresponding radiance maps for O IV 1401.2 Å (second column) and for O IV 1399.8 Å (sixth column). These radiance maps were computed by averaging the intensity within  $\pm 15 \text{ km s}^{-1}$  with respect to the core of each line. In addition, for both O IV lines, we computed maps of integrated emission in the blue wing at  $-50 \text{ km s}^{-1}$  (third and seventh column of the figure) and in the red wing at  $+50 \text{ km s}^{-1}$  (fourth and eighth column). Differences in the location of the emitting regions between the wings were determined from the composite maps (fifth and ninth column of the image). To ease the identification, all the panels show the contours that delimit the bulk of the surges visible in the Mg II  $k$  2796.3 Å radiance maps. The maps, especially the radiance ones of O IV 1399.8 Å, contain a horizontal strip of hot pixels; however, this does not affect our results.

Inspecting the image, we can see that both O IV 1401.2 Å and O IV 1399.8 Å radiance maps at the center of the line (second and sixth columns) look quite similar, although the surges appear to have a larger extension in O IV 1401.2 Å than in O IV 1399.8 Å, likely due to the fact that the former is a stronger line. The location of the brightest O IV regions can be found within the bulk of the surges and/or in their boundaries following the threads of the surges. This is particularly evident in the third and fourth rasters. The corresponding blue wing, red wing, and composite maps also provide pieces of evidence about the transition region counterpart of the surges in these weak O IV lines. For instance, in the third panel of the first row of the image, a black arrow indicates the location of enhanced O IV emission associated with finger-like threads of the surge. In addition, the maps also reveal a peculiar behavior in these threads: while several threads are observed collocated both in the blue and red wings of the O IV lines, some others show a significant shift in the position where they are detected in the blue and red wings of those lines. This asymmetry is similar to that noticed for Mg II  $k$  and Si IV 1402.8 Å lines by Guglielmino et al. (2019a) and it could



**Fig. 5.** Radiance maps showing the OIV emissivity of the surges. *From left to right:* radiance maps in the core of the Mg II  $k$  2796.3 Å line (*first column*); in the core of the O IV 1401.2 Å line (*second column*); in the blue and red wings at  $-50$  and  $50$  km s $^{-1}$  of O IV 1401.2 Å (*third and fourth columns*, respectively); and composite image of the blue and red wing radiance maps for O IV 1401.2 Å (*fifth column*). Equivalent maps are plotted for the O IV 1399.8 Å line (*sixth to ninth columns*). The contours in all the panels delimit the bulk of the surges visible in the Mg II  $k$  2796.3 Å radiance maps. The black arrow in the third panel of the first row indicates the location of enhanced OIV emission associated with finger-like threads of a surge. The OIV maps show a horizontal strip of hot pixels that does not affect our findings.

be evidence of helical and rotational motions typically associated with surges.

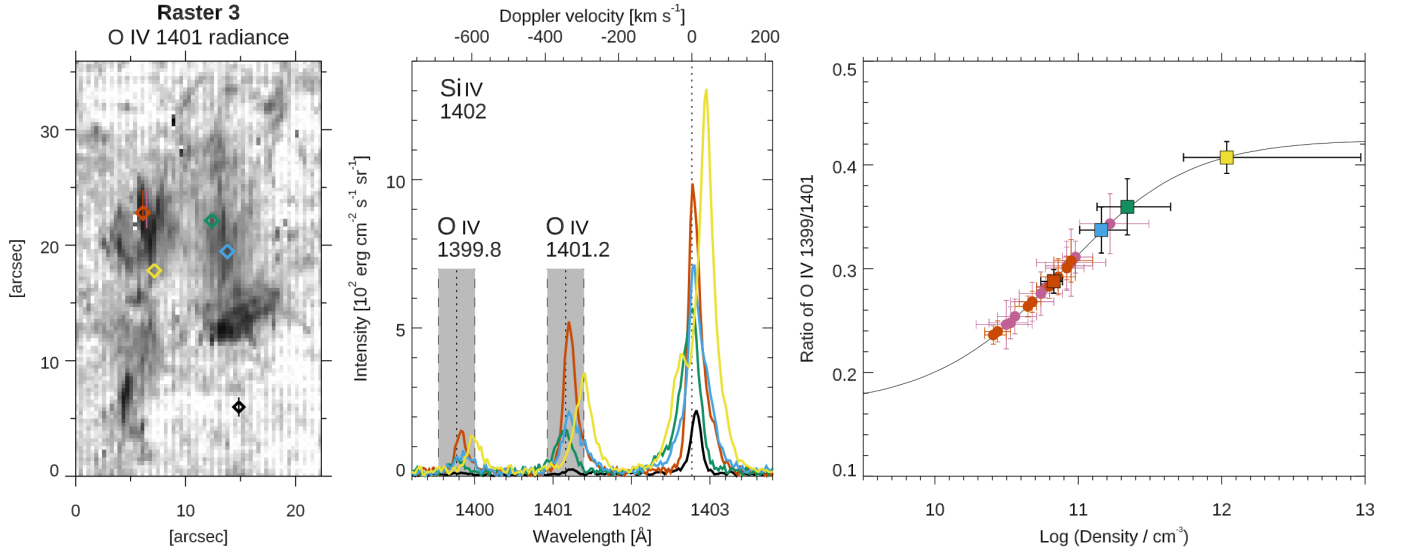
### 3.2.2. Spectral analysis and density diagnostics

To know how the OIV spectra related to the surges look, the left panel of Fig. 6 contains the O IV 1401.2 Å radiance map with a reversed color scale for the third raster. From our previous choice of locations in Fig. 2, we selected the pixels with a strong intensity in both O IV lines (orange, yellow, green, and cyan diamonds). In addition, we also considered some other pixels along a thread of the surge, approximately located at  $x = 6''$  and  $y \in [20'', 25'']$ , that have the greatest O IV intensity (orange and magenta segments). The black diamond in the map indicates a quiet-Sun area with a vertical segment of six pixels in which the spectrum has been averaged for comparison purposes. The corresponding surge and quiet-Sun spectra are shown in the middle panel of Fig. 6 using the same colors as for the diamonds. In this panel, we can see that the O IV intensity associated with the surge is a factor between 7 and 15 for the O IV 1399.8 Å line, and between 12 and 35 for the O IV 1401.2 Å line, which is brighter than the quiet-Sun average profile: surges have a transition

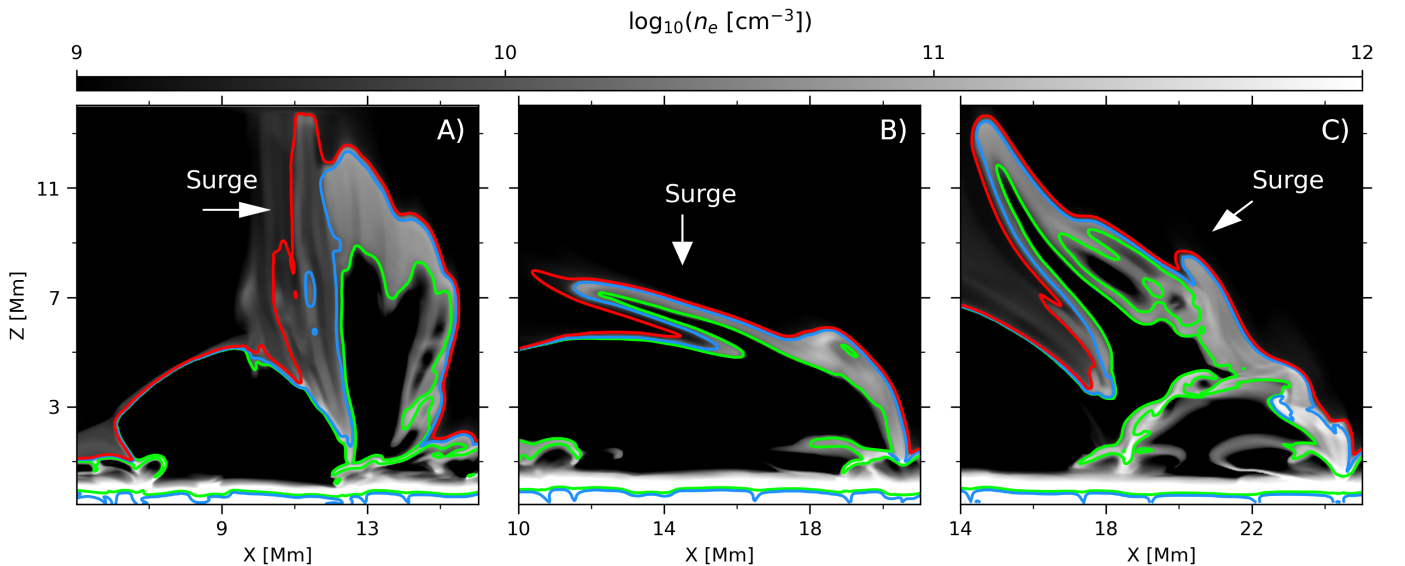
region counterpart even in the weakest FUV lines. This is moreover very different from UV bursts, where the O IV 1401.2 Å line is very weak in comparison to Si IV 1402.8, and the O IV 1399.8 Å line is absent (Peter et al. 2014; Young et al. 2018). With respect to the Si IV 1402.8 intensity, the ratio between the surge and the quiet-Sun region intensities ranges from 2.5 to 6.5, which agrees with the values previously found in IRIS observations (from 2 to 5) by Nóbrega-Siverio et al. (2017).

The simultaneous finding of O IV 1399.8 Å and 1401.2 Å allowed us to estimate the electron density in the transition region of the surges. Applying the density diagnostics method described in Sect. 2.2.3 for the selected pixels (diamonds and vertical segments in the left panel of Fig. 6), we obtained a density in the range of  $\approx 2.5 \times 10^{10} - 10^{12}$  cm $^{-3}$  as shown in the right panel of Fig. 5, where the yellow pixel lies at the high-density limit of the diagnostic method.

The next question to address is whether it is possible to establish a connection between the physical quantities that are derived for the transition region from the O IV analysis and the ones obtained for the mid- and low-chromosphere in Sect. 3.1.2 from the Mg II  $h&k$  lines. The green location, for instance, shows an O IV spectrum (middle panel of Fig. 6) that appears to have



**Fig. 6.** O IV spectral analysis and density diagnostics. *Left:* O IV 1401.2 Å radiance map (reversed color scale) for the third raster with superimposed colored diamonds (orange, green, blue, and yellow) and colored segments (orange and magenta) where the spectra are analyzed. For reference, the black diamond illustrates a quiet-Sun area with a vertical segment of six pixels in which the quiet-Sun spectrum has been averaged. *Middle:* FUV spectra in the Si IV 1402 Å passband for the positions indicated in the left map with diamonds. Gray shaded areas delimit the spectral regions at  $\pm 50 \text{ km s}^{-1}$  encompassing O IV 1399.8 Å and O IV 1401.2 Å lines. *Right panel:* results from the density diagnostics using the O IV 1399.8/1401.2 ratio for the surge pixels selected in the left panel.



**Fig. 7.** Electron number density for three different simulated surges. Contours of temperature are superimposed for  $T = 6 \text{ kK}$  (green),  $T = 10 \text{ kK}$  (blue), and  $T = 200 \text{ kK}$  (red). *Panel A:* surge from Nóbrega-Siverio et al. (2016) numerical experiment. *Panels B and C:* surges from the simulations by Nóbrega-Siverio et al. (2017, 2018).

been taken in a pixel almost at rest, while the electron density (right panel of Fig. 6) is around  $2 \times 10^{11} \text{ cm}^{-3}$ . For this location, the Mg II *h&k* inversions (Fig. 3) show that, at  $\log_{10}(\tau) = -6$ , the velocity is small ( $\sim 1 \pm 1 \text{ km s}^{-1}$ ), which is consistent with the Doppler shift found in O IV, and the density is close to  $10^{12} \text{ cm}^{-3}$  and decreases toward higher layers in the atmosphere, which also seems to be consistent if we assume that the O IV information comes from an upper and more rarefied layer. The orange pixel has an O IV spectrum exhibiting a slight redshift of about  $10 \text{ km s}^{-1}$  and a density of  $\approx 8 \times 10^{10} \text{ cm}^{-3}$ . This also appears to be in agreement with the results from the inversions, which show redshift velocities of  $\sim 11 \pm 3 \text{ km s}^{-1}$  and a density of  $5 \times 10^{11} \text{ cm}^{-3}$  that is rapidly decreasing upward. The yellow

and cyan locations show both redshifts in O IV; the former has larger velocities and more electron density. This general behavior is also found in the inversions, although the comparison is more complicated owing to the uncertainties involved in these two locations. In the following section, we compare these results with those from numerical simulations.

### 3.3. Comparison between observations and numerical experiments

To get a joint perspective between observations and theory, Fig. 7 contains electron number density maps for three different 2.5D numerical experiments at the time when surges are at

**Table 2.** Averaged electron number density in the temperature contours of the numerical experiments shown in Fig. 7.

$\log_{10}(n_e [\text{cm}^{-3}])$	Exp. A	Exp. B	Exp. C
At $T = 6 \text{ kK}$	10.6	10.8	10.5
At $T = 10 \text{ kK}$	11.4	12.7	12.5
At $T = 200 \text{ kK}$	9.6	9.3	9.5

their apex. The ejections are discernible as elongated structures and their locations are indicated with arrows. The three surges show finger-like structures similar to the observed ones. It is also possible to discern, especially in the left panel, that  $n_e$  shows variations within the surge that look like threads. Temperature contours are superimposed for  $T = 6 \text{ kK}$  (green),  $T = 10 \text{ kK}$  (blue), and  $T = 200 \text{ kK}$  (red), highlighting the multi-thermal structure of these phenomena. Focusing on the location of these contours, we can see that in some regions of the surge there are large temperature gradients with cool chromospheric plasma being close to the transition region plasma (where O IV emission originates), while in other regions, the transition region of the surge is far away (from a 100 kms up to a few megameters) from the cool core. This may explain why the location of the observed brightest O IV regions in Fig. 5 can be found within the bulk of the surges and/or in their boundaries following the threads of the surges. Around the  $T = 6 \text{ kK}$  contour, the surges have a clear core with a small electron number density. The number density increases from the core outward until reaching the transition region, where it decreases again. Table 2 contains the averaged  $n_e$  in the temperature contours for the three experiments as a zero order approximation to compare them with the observations. By doing this, we obtained number densities that are lower than those derived from observations. This means that, although qualitatively the experiments help us to understand the observed features, for a more accurate comparison we would need to include nonequilibrium effects to properly model the chromosphere and synthesize Mg II  $h&k$  lines in the experiments (see Sect. 4).

#### 4. Discussion and conclusions

In this paper, we have used observations from the IRIS satellite to study an episode of recurrent surges that appear to be associated with UV bursts in Active Region NOAA 12529 in April 2016. This is the first study of surges that characterizes their fundamental properties using the Mg II  $h&k$  as well as the O IV 1399.8 and 1401.2 Å spectral lines, combining to that end different methods such as the  $k$ -means algorithm, inversions with the STiC code, and density diagnostics. In addition, we employed numerical experiments using the Bifrost code to get a qualitative comparison between observations and theory. In the following, we summarize and discuss the main findings of this paper.

In Sect. 3.1, we have analyzed the near-UV spectra of surges through the Mg II  $h&k$  line, focusing on the properties of the surges as they have been mainly unexplored. In Sect. 3.1.1, we have used the  $k$ -means method in each of the four selected rasters to find Mg II  $h&k$  representative profiles in our observation and invert them. Thanks to this method, we reduced the number of profiles to invert by a factor 43.2, thus alleviating the computational and analysis effort, as well as being able to discern regions with peculiar profiles, such as in the case of surges and UV bursts. The  $k$ -means algorithm is not the only approach for this kind of study. For instance, Verma et al. (2021) employed

$t$ -distributed stochastic neighbor embedding ( $t$ -SNE), another machine learning algorithm, to classify the H $\alpha$  spectra with the aim of separating different regions to perform cloud model inversions. In Sect. 3.1.2, we have inverted our representative profiles from  $k$ -means finding that the surges have their most probable temperature around 6 kK, electronic number densities mostly concentrated from  $\sim 1.6 \times 10^{11}$  to  $10^{12} \text{ cm}^{-3}$ , and LOS velocities of a few  $\text{km s}^{-1}$ . In addition, we have also shown that the statistical distributions of these properties are very similar for the different surges, meaning that these ejections can be well constrained in terms of their physical quantities. The uncertainties' analysis shows that the values of temperature for the surges are very reliable for optical depths between  $\log_{10}(\tau) = -6.0$  and  $-3.2$ , the density values can only be considered between  $\log_{10}(\tau) = -6.0$  to  $-4.8$ , while the velocities are generally acceptable also between  $\log_{10}(\tau) = -6.0$  to  $-4.8$ , although it is preferable to examine them individually, pixel by pixel. These results have allowed us not only to gain new insights of fundamental properties of the surges, but also to constitute the first steps to characterize the surge spectral profiles in the main chromospheric lines in which they are observed. To the best of our knowledge, there is only one paper (Kejun et al. 1996) in which physical parameters of surges were computed through a two-component cloud model using H $\alpha$ , obtaining  $T = 8.5 \text{ kK}$  and  $n_e = 1.8 \times 10^{10} \text{ cm}^{-3}$ . More recently, inverting the Ca II 8542 Å line using the NLTE code NICOLE, Kuridze et al. (2021) obtained the electron density values of a large spicule off the limb. Their results show that  $n_e$  ranges from  $2.5 \times 10^{13} \text{ cm}^{-3}$ , at the spicule bottom, to  $\sim 2 \times 10^{10} \text{ cm}^{-3}$ , at its top. They argue that the density obtained is more akin to surges rather than typical spicules. In our case, the results from the inversion reveal a more constrained range for the electronic number density.

In Sect. 3.2, we have focused on the transition region counterpart of surges through the FUV spectra. Thanks to the relatively long exposure time used in the IRIS sequence (18 s for FUV, thus ensuring a good signal-to-noise ratio for faint lines in the transition region), in Sect. 3.2.1 we have been able to find emission in the O IV 1399.8 Å and 1401.2 Å lines related to the surges. This finding is relevant due to the following reasons:

(i) this is the first observational report of O IV 1399.8 Å and 1401.2 Å related to surges, in fact, previous FUV analysis could not detect any O IV emission because the exposure time was too short (Nóbrega-Siverio et al. 2017), or because the focus was on other lines (Guglielmino et al. 2019a);

(ii) it clearly demonstrates that surges have a transition region counterpart even in the weakest FUV lines (the oscillator strength, which is proportional to the line strength, of 1401.2 Å is  $5.1 \times 10^{-7}$  and the one for O IV 1399.8 Å is  $4.3 \times 10^{-7}$ , which is 16% smaller); and

(iii) it gives observational support to the theoretical findings by Nóbrega-Siverio et al. (2018), which show that the surges can have enhanced emissivity in lines such as O IV 1401.2 Å. We have also found that surges have threads with important blue- and redshifts that can appear to be spatially displaced (see composite maps of Fig. 5). This could be an indication of rotational motions in the transition region counterpart of the surges during their untwisting and ejection. A similar asymmetry in the locations of the blue- and redshifts within surges was also found in the Mg II  $k$  and Si IV 1402 Å lines by Guglielmino et al. (2019a), and they join the further indications of helical and rotational motions associated with surges described previously in the literature (see, e.g., Canfield et al. 1996; Jibben & Canfield 2004; Jiang et al. 2007; Bong et al. 2014). In Sect. 3.2.2, we

have selected the brightest O IV pixels in the surge to perform density diagnostics, finding that the electron number density ranges from, approximately,  $2.5 \times 10^{10}$  to  $10^{12} \text{ cm}^{-3}$ . The physical quantities obtained from O IV seem to be consistent with those derived from Mg II *h&k* inversions for these locations. However, one has to be cautious: the O IV intensity we observed can come from the integration along the LOS of different regions of the surge, as shown in the numerical experiments by Nóbrega-Siverio et al. (2018), so the inferred electron density could just be a weighted mean average electron density along the LOS. In addition, nonequilibrium ionization can be relevant for these emitting layers, and density diagnostics may be affected by an order of magnitude (Olluri et al. 2013). A possible way to better constrain the physical properties of surges in the future would be to use inversions combining the information from the chromosphere and transition region in a similar fashion as Vissers et al. (2019) have done for UV bursts and Ellerman bombs.

In Sect. 3.3, we qualitatively compare three simulations with the observations, finding similarities in terms of the topology that may also explain the location of the observed brightest O IV regions with respect to the bulk of the surges. Nóbrega-Siverio et al. (2016) show that the bulk of their simulated surge (population A in that paper) is quite concentrated around 6 kK, which agrees with the values from the inversions we have found here. The electron number density, instead, seems to be lower in the simulations. This could be related to the lack of nonequilibrium ionization in these experiments. Nóbrega-Siverio et al. (2020) demonstrate that the LTE assumption substantially underestimates the ionization fraction in most of the emerged region, which could lead to lower electron number densities in the subsequent surge. Therefore, numerical simulations of surges including the ionization out-of-equilibrium of hydrogen seems mandatory to match the physical properties derived from observations.

The combination of methods and results obtained in this paper open new possibilities for the analysis and diagnostics of surges. For instance, our inverted representative profiles and their corresponding representative model atmospheres will be included in the IRIS<sup>2</sup> (Sainz Dalda et al. 2019) database, improving its current capabilities to invert eruptive phenomena. These results may facilitate building automatic detection algorithms for these key ejections in observations. The simultaneous inversions of different spectral lines in which surges are observed will also allow us to compare and better constrain numerical experiments. In this vein, in the near future we expect to exploit the existing high-quality data by the Swedish 1-m Solar Telescope (SST) to perform a detailed characterization of surges in other lines such as Ca II *H&K* and Ca II 8542 Å (see, e.g., the publicly available co-aligned IRIS and SST data sets by Rouppe van der Voort et al. 2020). In addition, future observations from, for instance, the SPICE instrument on board the Solar Orbiter mission (Spice Consortium 2020) could help us to continue exploring the upper chromosphere and transition region of these important phenomena through lines such as H I 1025.7 Å,  $\log_{10}(T) = 4$  K; C III 977.0 Å,  $\log_{10}(T) = 4.5$  K; and O VI 1031.9 Å,  $\log_{10}(T) = 5.5$  K. In the same way, the 4-m diameter *Daniel K. Inoué* Solar Telescope (DKIST, Rimmele et al. 2020) could offer potential diagnostics at upper chromosphere and transition region temperatures of surges, for example, in the He I D3 and 10 830 Å lines.

*Acknowledgements.* This research is supported by the Research Council of Norway through its Centres of Excellence scheme, project number 262622, as well as through the Synergy Grant number 810218 (ERC-2018-SyG) of the

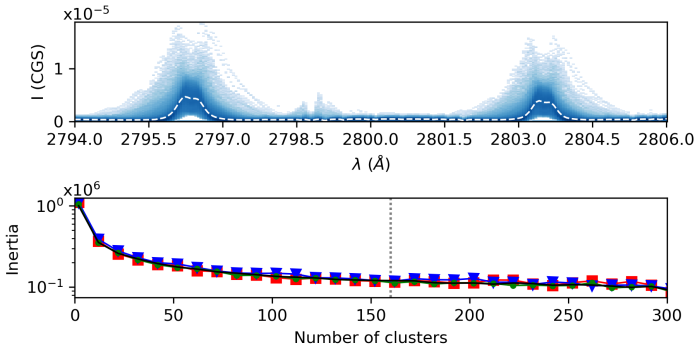
European Research Council, in addition to the project PGC2018-095832-B-I00 of the Spanish Ministry of Science, Innovation and Universities, and the ISSI Bern support for the team Unraveling surges: a joint perspective from numerical models, observations, and machine learning. This research has received funding from the European Commission's Seventh Framework Programme under the grant agreement no. 312495 (SOLARNET project) and from the European Union's Horizon 2020 research and innovation programme under the grant agreements no. 739500 (PRE-EST project) and no. 824135 (SOLARNET project). S. L. G. acknowledges support by the Italian MIUR-PRIN grant 2017APKP7T on Circumterrestrial Environment: Impact of Sun-Earth Interaction and support by the Italian Space Agency (ASI) under contract 2021-12-HH.0 to the co-financing INAF for the Italian contribution to the Solar-C EUVST preparatory science programme. A.S.D. contribution to this research is supported by NASA under contract NNG09FA40C (IRIS).

## References

- Asai, A., Ishii, T. T., & Kurokawa, H. 2001, *ApJ*, **555**, L65
- Barczynski, K., Harra, L., Kleint, L., Panos, B., & Brooks, D. H. 2021, *A&A*, **651**, A112
- Bong, S.-C., Cho, K.-S., & Yurchyshyn, V. 2014, *J. Korean Astron. Soc.*, **47**, 311
- Bose, S., Henriques, V. M. J., Joshi, J., & Rouppe van der Voort, L. 2019, *A&A*, **631**, L5
- Bose, S., Joshi, J., Henriques, V. M. J., & van der Voort, L. R. 2021a, *A&A*, **647**, A147
- Bose, S., Rouppe van der Voort, L., Joshi, J., et al. 2021b, *A&A*, **654**, A51
- Brooks, D. H., Kurokawa, H., & Berger, T. E. 2007, *ApJ*, **656**, 1197
- Canfield, R. C., Reardon, K. P., Leka, K. D., et al. 1996, *ApJ*, **464**, 1016
- Chae, J., Qiu, J., Wang, H., & Goode, P. R. 1999, *ApJ*, **513**, L75
- Chen, H. D., Jiang, Y. C., & Ma, S. L. 2008, *A&A*, **478**, 907
- Cho, K.-S., Cho, I.-H., Nakariakov, V. M., et al. 2019, *ApJ*, **877**, L1
- de la Cruz Rodríguez, J., Leenaarts, J., & Asensio Ramos, A. 2016, *ApJ*, **830**, L30
- de la Cruz Rodríguez, J., Leenaarts, J., Danilovic, S., & Uitenbroek, H. 2019, *A&A*, **623**, A74
- De Pontieu, B., Title, A. M., Lemen, J. R., et al. 2014, *Sol. Phys.*, **289**, 2733
- del Toro Iniesta, J. C., & Ruiz Cobo, B. 2016, *Liv. Rev. Sol. Phys.*, **13**, 4
- Del Zanna, G., Dere, K. P., Young, P. R., Landi, E., & Mason, H. E. 2015, *A&A*, **582**, A56
- Dudík, J., Del Zanna, G., Dzifčáková, E., Mason, H. E., & Golub, L. 2014, *ApJ*, **780**, L12
- Feldman, U., & Doschek, G. A. 1979, *A&A*, **79**, 357
- Fontenla, J. M., Avrett, E. H., & Loeser, R. 1993, *ApJ*, **406**, 319
- Gu, X. M., Lin, J., Li, K. J., et al. 1994, *A&A*, **282**, 240
- Gudiksen, B. V., Carlsson, M., Hansteen, V. H., et al. 2011, *A&A*, **531**, A154
- Guglielmino, S. L., Bellot Rubio, L. R., Zuccarello, F., et al. 2010, *ApJ*, **724**, 1083
- Guglielmino, S. L., Romano, P., & Zuccarello, F. 2017, *ApJ*, **846**, L16
- Guglielmino, S. L., Zuccarello, F., Young, P. R., Murabito, M., & Romano, P. 2018, *ApJ*, **856**, 127
- Guglielmino, S. L., Young, P. R., & Zuccarello, F. 2019a, *ApJ*, **871**, 82
- Guglielmino, S. L., Romano, P., Ruiz Cobo, B., Zuccarello, F., & Murabito, M. 2019b, *ApJ*, **880**, 34
- Hong, J., Ding, M. D., & Cao, W. 2017, *ApJ*, **838**, 101
- Huang, Z., Madjarska, M. S., Koleva, K., et al. 2014, *A&A*, **566**, A148
- Huang, Z., Madjarska, M. S., Scullion, E. M., et al. 2017, *MNRAS*, **464**, 1753
- Jiang, Y. C., Chen, H. D., Li, K. J., Shen, Y. D., & Yang, L. H. 2007, *A&A*, **469**, 331
- Jibben, P., & Canfield, R. C. 2004, *ApJ*, **610**, 1129
- Joshi, R., Chandra, R., Schmieder, B., et al. 2020a, *A&A*, **639**, A22
- Joshi, J., Rouppe van der Voort, L. H., & de la Cruz Rodríguez, J. 2020b, *A&A*, **641**, L5
- Joshi, R., Schmieder, B., Tei, A., et al. 2021, *A&A*, **645**, A80
- Kejun, L., Jun, L., Xiaoma, G., & Shuhua, Z. 1996, *Sol. Phys.*, **168**, 91
- Kim, Y.-H., Yurchyshyn, V., Bong, S.-C., et al. 2015, *ApJ*, **810**, 38
- Kirshner, R. P., & Noyes, R. W. 1971, *Sol. Phys.*, **20**, 428
- Kontogiannis, I., Tsiropoula, G., Tziotziou, K., et al. 2020, *A&A*, **633**, A67
- Kuckein, C., González Manrique, S. J., Kleint, L., & Asensio Ramos, A. 2020, *A&A*, **640**, A71
- Kuridze, D., Socas-Navarro, H., Koza, J., & Oliver, R. 2021, *ApJ*, **908**, 168
- Leenaarts, J., Pereira, T., & Uitenbroek, H. 2012, *A&A*, **543**, A109
- Leenaarts, J., Pereira, T. M. D., Carlsson, M., Uitenbroek, H., & De Pontieu, B. 2013, *ApJ*, **772**, 90
- Li, Z., Fang, C., Guo, Y., et al. 2016, *ApJ*, **826**, 217
- Li, H., Jiang, Y., Yang, J., et al. 2017, *ApJ*, **842**, L20
- Liu, Y., & Kurokawa, H. 2004, *ApJ*, **610**, 1136

- Liu, W., Berger, T. E., Title, A. M., & Tarbell, T. D. 2009, *ApJ*, **707**, L37
- MacTaggart, D., Guglielmino, S. L., Haynes, A. L., Simitev, R., & Zuccarello, F. 2015, *A&A*, **576**, A4
- Madjarska, M. S., Doyle, J. G., & de Pontieu, B. 2009, *ApJ*, **701**, 253
- McMath, R. R., & Mohler, O. 1948, *The Observatory*, **68**, 110
- Moreno-Insertis, F., & Galsgaard, K. 2013, *ApJ*, **771**, 20
- Moreno-Insertis, F., Galsgaard, K., & Ugarte-Urra, I. 2008, *ApJ*, **673**, L211
- Nelson, C. J., & Doyle, J. G. 2013, *A&A*, **560**, A31
- Nelson, C. J., Freij, N., Bennett, S., Erdélyi, R., & Mathioudakis, M. 2019, *ApJ*, **883**, 115
- Nishizuka, N., Shimizu, M., Nakamura, T., et al. 2008, *ApJ*, **683**, L83
- Nóbrega-Siverio, D., Moreno-Insertis, F., & Martínez-Sykora, J. 2016, *ApJ*, **822**, 18
- Nóbrega-Siverio, D., Martínez-Sykora, J., & Moreno-Insertis, F. 2017, & Rouppe van der Voort, L., *ApJ*, **850**, 153
- Nóbrega-Siverio, D., Moreno-Insertis, F., & Martínez-Sykora, J. 2018, *ApJ*, **858**, 8
- Nóbrega-Siverio, D., Moreno-Insertis, F., Martínez-Sykora, J., Carlsson, M., & Szydlarski, M. 2020, *A&A*, **633**, A66
- Olluri, K., Gudiksen, B. V., & Hansteen, V. H. 2013, *ApJ*, **767**, 43
- Ortiz, A., Hansteen, V. H., Nóbrega-Siverio, D., & van der Voort, L. R. 2020, *A&A*, **633**, A58
- Panos, B., Kleint, L., Huwylar, C., et al. 2018, *ApJ*, **861**, 62
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *J. Mach. Learn. Res.*, **12**, 2825
- Pesnell, W. D., Thompson, B. J., & Chamberlin, P. C. 2012, *Sol. Phys.*, **275**, 3
- Peter, H., Tian, H., Curdt, W., et al. 2014, *Science*, **346**, 1255726
- Pietarila, A., Socas-Navarro, H., & Bogdan, T. 2007, *ApJ*, **663**, 1386
- Polito, V., Del Zanna, G., Dudík, J., et al. 2016a, *A&A*, **594**, A64
- Polito, V., Reep, J. W., Reeves, K. K., et al. 2016b, *ApJ*, **816**, 89
- Rimmele, T. R., Warner, M., Keil, S. L., et al. 2020, *Sol. Phys.*, **295**, 172
- Robustini, C., Leenaarts, J., & de la Cruz Rodríguez, J. 2016, *A&A*, **590**, A57
- Robustini, C., Esteban Pozuelo, S., Leenaarts, J., & de la Cruz Rodríguez, J. 2019, *A&A*, **621**, A1
- Rouppe van der Voort, L., De Pontieu, B., Scharmer, G. B., et al. 2017, *ApJ*, **851**, L6
- Rouppe van der Voort, L. H. M., De Pontieu, B., Carlsson, M., et al. 2020, *A&A*, **641**, A146
- Roy, J.-R. 1973, *Sol. Phys.*, **32**, 139
- Ruan, G., Schmieder, B., Masson, S., et al. 2019, *ApJ*, **883**, 52
- Rutten, R. J., Vissers, G. J. M., Rouppe van der Voort, L. H. M., Sütterlin, P., & Vitas, N. 2013, *J. Phys. Conf. Ser.*, **440**
- Sainz Dalda, A., de la Cruz Rodríguez, J., De Pontieu, B., & Gošić, M. 2019, *ApJ*, **875**, L18
- Scherrer, P. H., Schou, J., Bush, R. I., et al. 2012, *Sol. Phys.*, **275**, 207
- Schmieder, B., Mein, P., Martres, M. J., & Tandberg-Hanssen, E. 1984, *Sol. Phys.*, **94**, 133
- Schrijver, C. J., & Higgins, P. A. 2015, *Sol. Phys.*, **290**, 2943
- Shimizu, T., Katsukawa, Y., Kubo, M., et al. 2009, *ApJ*, **696**, L66
- Spice Consortium (Anderson, M., et al.) 2020, *A&A*, **642**, A14
- Sterling, A. C., Moore, R. L., Falconer, D. A., et al. 2016, *ApJ*, **821**, 100
- Tandberg-Hanssen, E. 1959, *ApJ*, **130**, 202
- Tian, H., Yurchyshyn, V., Peter, H., et al. 2018, *ApJ*, **854**, 92
- Uddin, W., Schmieder, B., Chandra, R., et al. 2012, *ApJ*, **752**, 70
- Vargas Domínguez, S., Kosovichev, A., & Yurchyshyn, V. 2014, *ApJ*, **794**, 140
- Verma, M., Denker, C., Diercke, A., et al. 2020, *A&A*, **639**, A19
- Verma, M., Matijevič, G., Denker, C., et al. 2021, *ApJ*, **907**, 54
- Vissers, G. J. M., Rouppe van der Voort, L. H. M., & Rutten, R. J. 2013, *ApJ*, **774**, 32
- Vissers, G. J. M., de la Cruz Rodríguez, J., Libbrecht, T., et al. 2019, *A&A*, **627**, A101
- Viticchié, B., & Sánchez Almeida, J. 2011, *A&A*, **530**, A14
- Wang, H., & Liu, C. 2012, *ApJ*, **760**, 101
- Wang, J. F., Zhou, T. H., & Ji, H. S. 2014, *Chin. Astron. Astrophys.*, **38**, 65
- Watanabe, H., Vissers, G., Kitai, R., Rouppe van der Voort, L., & Rutten, R. J. 2011, *ApJ*, **736**, 71
- Yang, H., Chae, J., Lim, E.-K., et al. 2013, *Sol. Phys.*, **288**, 39
- Yang, H., Chae, J., Lim, E.-K., et al. 2014, *ApJ*, **790**, L4
- Yang, L., Peter, H., He, J., et al. 2018, *ApJ*, **852**, 16
- Yokoyama, T., & Shibata, K. 1996, *PASJ*, **48**, 353
- Young, P. R., Tian, H., Peter, H., et al. 2018, *Space Sci. Rev.*, **214**, 120
- Zhang, J., Wang, J., & Liu, Y. 2000, *A&A*, **361**, 759
- Zhang, Q. M., & Ji, H. S. 2014, *A&A*, **561**, A134

## Appendix A: Processing for the $k$ -means algorithm



**Fig. A.1.** Data information for the  $k$ -means clustering. *Top*: Joint-PDF of all the calibrated Mg II h&k line profiles in CGS units ( $\text{erg cm}^{-2} \text{s}^{-1} \text{sr}^{-1} \text{Hz}^{-1}$ ) within the sub-FoV of Fig. 1 for the four rasters. The corresponding average profile is shown as a dashed line. *Bottom*: Behavior of the inertia of the  $k$ -means method with increasing clusters to choose a suitable number through the elbow technique for raster 1 (red square), raster 2 (blue triangle), raster 3 (green circle), and raster 4 (black line). Our cluster choice for all the rasters,  $k = 160$ , is indicated by the vertical line.

To obtain meaningful results out of the  $k$ -means clustering, it is necessary to process the observational data properly. To that end, we have carried out the following steps.

1. Selection of the domain. We chose the sub-FoV of Fig. 1 for our  $k$ -means clustering, thus reducing the number of profiles to be grouped.
2. Removing bad data. Any profile affected by data loss or cosmic rays was replaced by zero values. This is helpful as a benchmark: our data have 327 profiles that are affected and  $k$ -means grouped them together in one single cluster, thus being able to easily mask them in the figures of the paper.
3. Radiometric calibration. We performed radiometric calibration<sup>3</sup> of the observational data to convert them into physical units. This step is essential as it allows us to use the results obtained from the  $k$ -means method as input for the inversions.
4. Cropping the spectra. We cropped the spectra to the range between 2794 to 2806 Å, which was sampled with 237 points. This is enough to contain the whole Mg II h&k line. The top panel of Fig. A.1 shows the distribution of all the calibrated profiles through a joint-PDF plot as well as the average profile (dashed line).
5. Standardization of individual features. This step is a common requirement for machine learning estimators to avoid bad behavior if the features are not more or less normally distributed. Our features are the wavelength positions, so for all the 237 positions between 2794 to 2806 Å, we carried out a standardization of the intensity.

After processing all the data, we performed the  $k$ -means method using the tools available on Scikit-learn (Pedregosa et al. 2011) and selecting 160 clusters for each of the four rasters. The initialization of clusters was repeated 100 times, keeping the best output of the consecutive runs in terms of inertia. We conclude that  $k = 160$  is an appropriate number of clusters through three different methods: visual comparison of the representative profiles with the observed profiles belonging to a given cluster through joint-PDFs; using the benchmark of the bad pixels mentioned before; and checking the behavior of the inertia with the number of clusters (the elbow technique shown in the bottom panel of Fig. A.1). In the figure, the regime  $k \geq 80$  shows little variation of the inertia for the four rasters. Since our purpose pertains to inversions and not finding the minimum number of clusters in which the data can be grouped, we can take any  $k$  in this regime because the inversions give very similar results.

<sup>3</sup> See the documentation about IRIS Level 2 data calibration in [https://iris.lmsal.com/iris2/iris2\\_chapter04\\_01.html#calibration-of-iris-level-2-data](https://iris.lmsal.com/iris2/iris2_chapter04_01.html#calibration-of-iris-level-2-data)