Large-Scale Spatial Cross-Calibration of *Hinode*/SOT-SP and *SDO*/HMI

David F. Fouhey [ID],[1] Richard E. L. Higgins [ID],[1] Spiro K. Antiochos [ID],[2] Graham Barnes [ID],[3] Marc L. DeRosa [ID],[4] J. Todd Hoeksema [ID],[5] K. D. Leka [ID],[3] Yang Liu [ID],[5] Peter W. Schuck [ID],[6] and Tamas I. Gombosi [ID][7]

[1]*University of Michigan, Department of Electrical Engineering and Computer Science, Ann Arbor, MI*
[2]*University of Michigan, Department of Climate and Space Sciences and Engineering, Ann Arbor, MI*
[3]*NorthWest Research Associates, Boulder, CO*
[4]*Lockheed Martin Solar and Astrophysics Laboratory, Palo Alto, CA, USA*
[5]*Stanford University, Stanford, CA*
[6]*NASA GSFC, Silver Spring, MD*
[7]*University of Michigan, Department of Climate and Space, Center for Space Environment Modelling, Ann Arbor, MI*

ABSTRACT

We investigate the cross-calibration of the Hinode/SOT-SP and SDO/HMI instrument meta-data, specifically the correspondence of the scaling and pointing information. Accurate calibration of these datasets gives the correspondence needed by inter-instrument studies and learning-based magnetogram systems, and is required for physically-meaningful photospheric magnetic field vectors. We approach the problem by robustly fitting geometric models on correspondences between images from each instrument's pipeline. This technique is common in computer vision, but several critical details are required when using scanning slit spectrograph data like Hinode/SOT-SP. We apply this technique to data spanning a decade of the Hinode mission. Our results suggest corrections to the published Level 2 Hinode/SOT-SP data. First, an analysis on approximately 2,700 scans suggests that the reported pixel size in Hinode/SOT-SP Level 2 data is incorrect by around 1%. Second, analysis of over 12,000 scans show that the pointing information is often incorrect by dozens of arcseconds with a strong bias. Regression of these corrections indicates that thermal effects have caused secular and cyclic drift in Hinode/SOT-SP pointing data over its mission. We offer two solutions. First, direct co-alignment with SDO/HMI data via our procedure can improve alignments for many Hinode/SOT-SP scans. Second, since the pointing errors are predictable, simple post-hoc corrections can substantially improve the pointing. We conclude by illustrating the impact of this updated calibration on derived physical data products needed for research and interpretation. Among other things, our results suggest that the pointing errors induce a hemispheric bias in estimates of radial current density.

1. INTRODUCTION

The Sun's photosphere is co-observed by a variety of instruments that provide tradeoffs between design considerations that range from spectral to spatial sampling. Each instrument independently provides a view of the magnetic field, and by combining multiple instruments with varying capabilities, one can capture phenomena from different vantage points, obtain information that is impossible to extract from one instrument alone, and calibrate the instruments. For instance, Sainz Dalda (2017) use co-aligned data to better understand the magnetograms produced by the pipelines of *Hinode*/SOT-SP and *SDO*/HMI. Similarly, our previous work, Higgins et al. (2022) shows how to use co-aligned *Hinode*/SOT-SP and *SDO*/HMI data to produce synthetic magnetograms from *SDO*/HMI Stokes vectors with characteristics similar to *Hinode*/SOT-SP

data. These efforts on using multiple instruments depend critically on precise co-alignment of the data.

In this paper, we demonstrate and analyze the spatial co-alignment of the *Solar Optical Telescope-Spectro-Polarimeter* (SOT-SP; Tsuneta et al. 2008) on the *Hinode* satellite (Kosugi et al. 2007) with the *Helioseismic and Magnetic Imager* (HMI; Schou et al. 2012) on the *Solar Dynamics Observatory* (SDO; Pesnell et al. 2012). Since the focus of our work is magnetograms, we specifically focus on Level 2 *Hinode*/SOT-SP data (Community Spectropolarimetric Analysis Center (CSAC) 2006) inverted by MERLIN (Lites et al. 2006) and for *SDO*/HMI, the `hmi.ME_720s_fd10` series inverted by the *SDO*/HMI pipeline version of VFISV (Borrero et al. 2011; Centeno et al. 2014).

We obtain the relative scaling and pointing of the two instruments using techniques that are common in

computer vision. We approach the problem via a standard computer vision practice of fitting geometric models (see e.g., the books of Hartley & Zisserman 2004 and Szeliski (2010)) to correspondences automatically extracted from image data. These correspondences come from the point-based features SIFT (Lowe 2004) and ORB (Rublee et al. 2011) that are extracted at distinctive and easily localizable locations in the image. Because the correspondences are extracted by matching texture in small image regions, they are likely contaminated with outliers. However, by fitting models robustly using RANSAC (Fischler & Bolles 1981) to reject outliers, we are able to characterize accurately the relationship between the instruments. This general framework of fitting models to correspondences is well-understood and validated in computer vision and has been applied successfully to tasks ranging from automatic panorama construction (Brown & Lowe 2003) to building precise 3D reconstructions from photo collections (Snavely et al. 2007; Schönberger & Frahm 2016) to performing simultaneous localization and mapping (Mur-Artal & Tardós 2017), and far more. However, several key adjustments to this standard procedure are needed for the *Hinode*/SOT-SP data because the data comes from a scanning slit spectrograph. This spectrograph captures the Sun at different times of evolution and does not uniformly sample the scan spatially. Both challenges can be overcome by modifications to this alignment procedure that we describe in Section 3.

We apply this technique to thousands of *Hinode*/SOT-SP scans that span nearly a decade and come from both of *Hinode*/SOT-SP's fast and normal modes (Tsuneta et al. 2008; Lites et al. 2013), which acquire data with different effective pixel sizes (described in further detail in Section 2). By aligning the data, we can extract parameters that ought to not vary between scans, namely the relative scaling of the instruments, as well as parameters that will vary between scans, namely the relative pointing information and instrument rotation. We can align the data while letting the relative instrument scaling vary freely. By performing this procedure across multiple scans, we can obtain an updated empirical estimate of the relative scaling of the instruments. Once a new scale has been fit, we can estimate updated pointing information and compare with the reported pointing information.

The first key result of this work, described in Section 4 is a suggested correction to the reported *Hinode*/SOT-SP Level 2 data pixel size. Fit scale parameters suggest that data are consistently explained better with different scaling parameters than is reported. This scale correction is nearly the same for both fast and normal scans

and is near constant across observations spanning nearly a decade of *Hinode*/SOT-SP's missions. The fit scale additionally shows no strong correlation with spatial location or time (represented at multiple frequencies).

The second key result is a suggested update to *Hinode*/SOT-SP's pointing information, described in Section 5. This information is known to be inaccurate (Hurlburt et al. 2009, Appendix C) and our approach provides highly accurate correspondence between the scans, providing a way to update the pointing information reported with the *Hinode*/SOT-SP scan. These pointing updates are typically on the order of dozens of arcseconds and show a strong directional bias. Examination of the pointing updates as a function of time and temperature readings suggest that these residuals can be well-explained by secular trends in pointing information during the Hinode mission as well as yearly thermal changes during eclipse season. Co-alignment with *SOHO*/MDI (Scherrer et al. 1995) further suggests that the pointing error was substantially smaller earlier in the mission. New *Hinode*/SOT-SP data can be, of course, co-registered to new *SDO*/HMI data to provide improved pointing information. If this registration fails or *SDO*/HMI is no longer available, our results show that the drift is relatively predictable and one can obtain substantially more accurate pointing information (about an order of magnitude smaller error) via a simple predictive model.

Section 6 describes the physical implications of the updated pointing information. The change in pointing information affects both the heliographic coordinates of each pixel as well as distance between the parts of the sun depicted by each pixel. These updates lead to distinctive spatial patterns and, due to the bias in the pointing update, hemispheric bias in quantities such as the total electric current. A summary discussion and concluding remarks are presented in Section 7.

2. DATA

We work with two sources of data, namely the pipelines of *Hinode*/SOT-SP and *SDO*/HMI, focusing on 11 years of *Hinode*/SOT-SP Level 2 data from 1 January 2011 until 31 December 2021 as well as *SDO*/HMI data that was observed at the same time. The goal of this section is to reintroduce salient aspects of the data acquisition pipelines that are needed to understand the technical conclusions of the paper. We begin by reprising the data pipeline, consisting of an overview of the acquisition process (Section 2.1) and a description of how scale and pointing information changes in the *Hinode*/SOT-SP and *SDO*/HMI data pipelines (Section 2.2. A more full description of the characteristics of

the data products is beyond the scope of the paper. We refer the interested reader to Tsuneta et al. (2008) and Hoeksema et al. (2014) for descriptions of the data products of *Hinode*/SOT-SP and *SDO*/HMI respectively. We then discuss the specific data used in this paper. First, we introduce the data filtering process (Section 2.3) and then the datasets that we use (Section 2.4).

2.1. *Overview of Data Acquisition Pipelines*

SDO/HMI is an imager that prioritizes the acquisition of *spatially* coherent observations. While we show images with multiple channels from *SDO*/HMI, its underlying observation is a single 4096×4096 full-disk image of the full sun. Over time, multiple such images with different passbands and polarization states are acquired and analyzed in order to recover the Stokes vector. The Stokes vector observation at each pixel is then inverted to produce a magnetogram that is recorded at a nominal time.

Hinode/SOT-SP, on the other hand, is a scanning-slit spectrograph that prioritizes spectrally coherent observations. While we show data from *Hinode*/SOT-SP throughout the paper as images, its underlying observation is a single vertical scanline with dense spectral sampling of the Stokes vector. These scanlines form the basis of *Hinode*/SOT-SP's Level 1 data. In the data pipeline, a number of these scanlines are concatenated horizontally and the Stokes vector at each pixel in each scanline is inverted. By placing W scanlines side-by-side, this produces a $H \times W$ magnetogram that is reported as the *Hinode*/SOT-SP Level 2 data (Community Spectropolarimetric Analysis Center (CSAC) 2006). For the data used in this study, the acquisition of all the scanlines in a magnetogram takes roughly 30 minutes to an hour. These magnetograms are not acquired with a regular cadence, since *Hinode*/SOT-SP has targets of interest selected and then acquired. Given the difference in data acquisition times, for any given *Hinode*/SOT-SP scan, there are likely multiple contemporaneous *SDO*/HMI scans.

The *Hinode*/SOT-SP acquisition process means that the resulting magnetogram is not at all like *SDO*/HMI's image: each x coordinate in a *Hinode*/SOT-SP magnetogram represents a scan *index*, rather than a scan *location*. During acquisition, the spectrograph makes observations at slits, with its position controlled by a tracker. Except in rare cases where the observed slit positions are perfectly evenly spaced, the pattern of the observations is *not* uniform but instead contains jumps where the data from individual slit positions are corrupt or otherwise unusable. Since the position of these jumps is known and recorded in the `Mechanical_Slit_Position`

field, the $H \times W$ magnetogram can instead be thought of as shorthand for a wider, $H \times W'$ image (with $W' \geq W$) that has columns missing. While the wider $H \times W'$ image has missing data, it *can* be treated as an image where the x coordinate represents a location just like the y coordinate.

Throughout the paper, we use the word *scan*. In the case of *SDO*/HMI, this refers to a single vector magnetogram from *SDO*/HMI at a given nominal timestamp. For *Hinode*/SOT-SP, this refers to an accumulation of scanlines into a single $H \times W$ vector magnetogram, also with a nominal timestamp. We stress, however, that the magnetogram cannot be treated as an image and that throughout, we use the information about the slit positions to account for non-uniform scans. In the code release for the paper (to be provided before publication), we will provide demonstration code that does this.

2.2. *Scales and Pointing Throughout the Hinode/SOT-SP and SDO/HMI Pipelines*

Due to the flexibility of the *Hinode*/SOT-SP instrument, the pixels of its magnetograms come in a variety of scales. We focus on two types that are roughly square, namely *normal* mode and *fast* mode. In normal mode, the *Hinode*/SOT-SP Level 2 data headers report a pixel size of $0.15999''$ for `YSCALE` (the size of the pixel in y) and $0.1486''$ for `XSCALE` (the size of the pixel in x). Fast mode scans have double this pixel size due to the ability of the instrument to combine data from neighboring pixels prior to their transmission back to Earth, and due to their smaller telemetry requirements, fast scans are far more common than normal scans in the date range we study. There are other *Hinode*/SOT-SP scans whose pixels have even more non-square aspect ratios; we ignore these in our study because they are substantially less frequent.

While the *Hinode*/SOT-SP Level 2 data reports pixel sizes of $0.15999''$ and $0.1486''$, this is not the case throughout the pipeline. The Level 1 data reports that `CDELT2` (equivalent to `YSCALE`) is $0.1585''$ and the nominal scanning step size (roughly equivalent to `XSCALE`) is $0.1476''$. The reason for the change is an earlier cross-instrument spatial calibration study from Centeno et al. (2009) that calibrated a continuum image synthesized from *Hinode*/SOT-SP and the G band from *Hinode*/BFI. Our results contradict this change and suggest that *Hinode*/SOT-SP's Level 1's reported scale is likely correct. We discuss this contradiction in more detail in Section 7.

Because *Hinode*/SOT-SP does not observe the full disk, its pointing information is obtained via the *Hinode*'s attitude and orbit control system (AOCS) (Kosugi

et al. 2007). The AOCS obtains its pointing information from sun sensors and star trackers. This information is used directly to calculate XCEN and YCEN via the SC_ATTX and SC_ATTY keywords and calibration constants. As discussed in Hurlburt et al. (2009), this pointing information is known to be inaccurate.

In comparison, *SDO*/HMI is substantially easier to calibrate because it can observe the full disk of the Sun. Its estimated CDELT1/2 varies slightly over the period that we study with values ranging from $0.50400''$ to $0.50439''$. For simplicity, throughout the analysis we use a nominal scale of $0.50428''$, the median reported CDELT1/2 of the HMI scans used in this analysis. This variance has no consequence on the results since the range of scales amounts to a variation of 0.076% of the median value. An earlier version of this study used a nominal CDELT1/2 closer to the bottom range of *SDO*/HMI's scale and produced numbers that were generally the same to three significant figures.

2.3. *Hinode/SOT-SP Scan Type Identification*

The goal of our analysis is to investigate the co-alignment between *Hinode*/SOT-SP and *SDO*/HMI, and so we automatically identify a subset of the *Hinode*/SOT-SP scans where this is most likely to succeed. We select scans that: (1) have near-square pixel aspect ratio; (2) have monotonic and relatively smooth spectrograph slit position during capture; and (3) are likely on disk as opposed to showing the poles. We next describe how we operationalize these definitions.

We work with normal (XSCALE = $0.1486''$, YSCALE = $0.1599''$) and fast (XSCALE = $0.297''$, YSCALE = $0.320''$) scans, which we identify by comparing reported XSCALE and YSCALE with fixed values.

We select for scans that are sufficiently smooth to function closely to images. While we introduce techniques that handle non-uniform sampling due to jumps in slit position, scans that substantially deviate from a uniform scanning pattern are likely to introduce noise. Therefore we reject scans based on two criteria using the Mechanical_Slit_Position data. First, if slit position does not monotonically increase, we reject the scan. The slit position may not monotonically increase if, for instance, *Hinode*/SOT-SP repeatedly scans an active region in a looping pattern. These are insufficiently frequent to merit a special procedure to further split these into images. Second, we compute the increase in slit position from index to index and compare it to a typical increase. Slit position increases that are $10\times$ or more than the median slit increase are identified as discontinuities, and we reject scans with a discontinuity in the

middle 96% of the slit. These have large jumps in slit position, and our unreliable for registration.

Finally, we automatically identify polar scans by comparing the top and bottom rows of the image with the median continuum intensity. We remove polar scans since they tend to have too few features to be reliably aligned with our proposed methods. If more than the majority of pixels in five or more rows have continuum intensity (as recorded by Original_Continuum_Intensity) less than half the median continuum intensity value, we identify the image as polar.

2.4. *Datasets*

We use a large dataset of *Hinode*/SOT-SP scans as well as two subsets of this dataset. Pilot studies with smaller subsets of data identified that pointing was highly varied but scale was more or less constant. Given that pointing was highly varied, we used the full dataset in order to get a full picture of the pointing information. On the other hand, since scale did not vary over time and space and the validation done was computationally expensive, we present analysis of scale only on the subset.

Our full dataset, which we refer to as the *Full Dataset*, is all the scans from 1 January 2011 through 31 December 2021. This dataset comprises 23,430 scans, of which 16,564 meet our criterion, and 12,062 can be co-aligned using our methods. These 12,062 scans are aligned with 50,943 *SDO*/HMI observations (amounting to over 40TB of *SDO*/HMI data). We use two large random subsets of the main dataset for analyzing the relative scale of the two instruments. Our *Fast Subset* was obtained by: selecting a random 1500 fast scan subset of the full dataset, of which 988 met our criterion, and of which 760 scans can be co-aligned with our methods. Our *Normal Subset* was obtained by taking all the normal scans in the dataset that met our criterion, of which 503 scans can be co-aligned with our methods.

3. ALIGNMENT

Our procedure seeks to find a parametric model with a small number of parameters that co-aligns features automatically found in and matched between *Hinode*/SOT-SP and *SDO*/HMI data. Some of the model parameters, such as the relative scaling of each camera's pixels, do not change between scans. Others, such as the relative translation and rotation, change between the scans. These transformations are fit on point correspondences with a combination of non-linear least-squares (Marquardt 1963) and a robust fitting technique named RANSAC (Fischler & Bolles 1981). Once fit, one
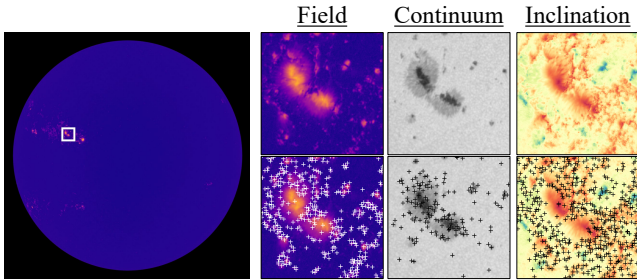
Figure 1. Locations of SIFT features found on *SDO*/**HMI data.** ORB features are found at similar distinctive regions in the image, but slightly different due to differences in the precise detection scheme In both cases, the features can then be matched to features found on *Hinode*/SOT-SP data to generate correspondences between the two datasets. These correspondences contain both outliers and noise, but the fitting methods used are robust to both.

can examine the parameters that were fit to data to both estimate the relative scaling of the instruments' observations as well as estimate the pointing information.

3.1. *Features and Feature Correspondence*

We fit the models on SIFT (Lowe 2004) and ORB (Rublee et al. 2011) features extracted on similar data from both instruments. These features are commonly used in photogrammetry applications for high-precision measurements, however a full description is beyond the scope of the work. Briefly, given two images $\mathbf{I}$ and $\mathbf{I'}$, each correspondence pipeline can generate correspondences between two images $[x, y] \leftrightarrow [x', y']$ such that point $[x, y]$ in image $\mathbf{I}$ and point $[x', y']$ in image $\mathbf{I'}$ both depict the same phenomena. Colloquially, image $\mathbf{I}$ at $[x, y]$ is said to "match" image $\mathbf{I'}$ at $[x', y']$. These correspondences are based on local intensity information in the image, extracted at distinctive parts of the image (often found via an approximation of the Laplacian of Gaussian filter). However, since the features are matched based on local intensity, there are many *outliers*, or correspondences between locations that look similar but are not the same.

One challenge is that the *Hinode*/SOT-SP is a scanning slit spectrograph that takes tens of minutes to finish a spatial scan. During this time, multiple complete sets of spectral- and polarimetric images are acquired by *SDO*/HMI. Given a set of correspondences $[x_i, y_i] \leftrightarrow [x'_i, y'_i]$ between a single *Hinode*/SOT-SP scan and single *SDO*/HMI scan taken at time t_{hmi}, one can further filter correspondences based on time. Each x coordinate in a *Hinode*/SOT-SP scan is associated with a timestamp of the scan $t(x)$ given by the `Times` field. One can simply remove correspondences $[x_i, y_i] \leftrightarrow [x'_i, y'_i]$

where $|t(x_i) - t_{\mathrm{hmi}}| > \delta$ where δ is an empirically tunable parameter set. We set δ to 24 minutes to strike a balance between two competing goals. If δ is too small, the correspondences used for transformation will come from a small time window and therefore come from a small number of scanlines. Estimating the scale in x with a few scanlines will be unreliable since the noise in localization of the points of interest will be a substantial fraction of the range of the data. On the other hand, while making δ bigger ensures that localization noise is a small fraction of the range of the data, making δ too big will result in correspondences coming from substantially different observations.

Practically speaking, this means that given a *Hinode*/SOT-SP scan and a *SDO*/HMI scan from nearly the same time, one can automatically generate correspondences between *SDO*/HMI data and parts of the *Hinode*/SOT-SP scan that were taken within a time window near the *SDO*/HMI scan.

3.2. *Data and Data Preprocessing*

One strong advantage of the correspondence-based approach is that the quantities that are matched do not have to be precisely the same. In fact, the methods used were designed to be robust to substantial nonlinear illumination changes as well as changes in viewing angle (Mikolajczyk & Schmid 2005). Mathematically, this stems from the fact that SIFT features are based on gradient orientation angles that have been spatially-histogrammed. The use of orientation *angle* rather than orientation *value* gives invariance to the scaling of the quantities, and the use of a spatial histogram gives invariance to small pixel perturbations. Similarly, ORB features are based on relative intensity comparisons after smoothing, which give similar invariance to scaling and perturbations.

This flexiblity to extract correspondences that are related (e.g., by an unknown nonlinear transformation) but not identical is appealing for cross-instrument analysis. Some quantities that are good for generating correspondence refer to different quantities: the *Hinode*/SOT-SP field represents field strength while the field reported by *SDO*/HMI is the flux density with an assumed magnetic fill factor of unity. Even when data refer to the same physical quantity, there can be differing characteristics of the pipeline outputs: as Sainz Dalda (2017) and Higgins et al. (2022) illustrate, *SDO*/HMI and *Hinode*/SOT-SP inclination angles tend to be farther from the imaging plane (90°) in plage regions compared to *SDO*/HMI. These differences arise from differing instrument characteristics (e.g., spectral sampling) as well as treatment of key Milne-Eddington model pa-
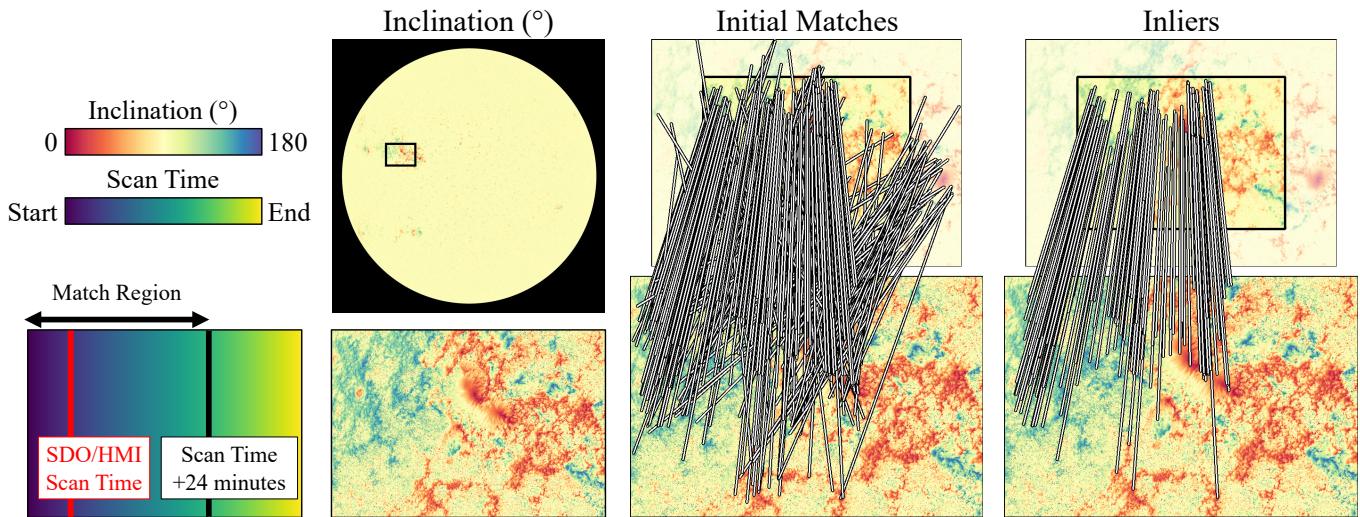
Figure 2. Diagram of Matching and Transformation Fitting. Given a Level 2 *Hinode*/SOT-SP scan and *SDO*/HMI scan, one can identify the contemporaneous part of the *Hinode*/SOT-SP scan and filter to the correspondences on contemporaneous *Hinode*/SOT-SP pixels. For clarity, we show only a subset of the matches that originate near the final fit. These have many outliers. After robust fitting, there are a far smaller number of inliers that are well-described by a single model. For clarity, the diagram shows the Level 2 *Hinode*/SOT-SP scan as it is stored, a set of columns next to each other. In practice, each correspondence's x coordinate is mapped to the slit position to account for the possibility of non-uniform spacing.

rameters (e.g., whether a magnetic fill fraction α is fit as in *Hinode*/SOT-SP or assumed as *SDO*/HMI) (Leka et al. 2022).

We generate correspondences on multiple pairs of data from *Hinode*/SOT-SP Level 2 and *SDO*/HMI `hmi.ME_720s_fd10` data. The first pair of data we match is the reported field strengths of the pipelines (`Field_Strength` for *Hinode*/SOT-SP and `field` for *SDO*/HMI). Note again, that *SDO*/HMI's field is flux density with an assumed unity fill fraction. The second and third are inclination, (`Field_Inclination` in *Hinode*/SOT-SP and `inclination` in *SDO*/HMI) and azimuth (`Field_Azimuth` in *Hinode*/SOT-SP and `azimuth` in *SDO*/HMI), after flipping the HMI data to acocunt for differences in `CROTA2`. Finally, we generate correspondences from continuum intensity (`Original Continuum Intensity` and `continuum` from `hmi.ic_720s`); polarization and field strength (`Polarization` and `field`); and finally Stokes V/circular polarization, where *Hinode*/SOT-SP's `StokesV_Magnitude` is compared with each of the six V filtergrams from *SDO*/HMI's `hmi.s_720s` series.

The feature methods that we use (SIFT, ORB) were designed for consumer image formats, and so operate on grayscale unsigned 8-bit images. We linearly remap the data to `uint8` by mapping each value x to $255(x - v_{min})/(v_{max} - v_{min})$ and clipping to between 0 and 255. We set the linear range with the goal of ensuring the 255 values are used on the bulk of the distribution, and

set v_{min} and v_{max} to: 0 and 2000 for field strength (to ensure that plage has sufficient contrast); and 20 and 160 for inclination (to avoid spending a large fraction of the range on rare values); and the 1st and 99th percentile of the data (excluding NaNs) otherwise (again, to avoid using much of the range on rare values).

3.3. *Fitting Image Transformation Models*

We fit a series of transformations that describe the relationship between the two pieces of data. Once fit, each transformation can be expressed as an affine transformation from *Hinode*/SOT-SP pixel coordiantes to *SDO*/HMI pixel coordinates, or $f([x, y]) = [x', y'] = \mathbf{A}[x, y] + [t_x, t_y]$ although each model puts constraints on the form of $\mathbf{A}$. This affine transformation can, of course, be inverted to enable inverse warping, i.e., to enable estimating the *Hinode*/SOT-SP values at each *SDO*/HMI pixel.

While one could identify the translation between the scans using cross-correlation, fitting a transformation on a set of point correspondences offers substantial advantages over cross-correlation. The most important advantage is that correspondences enable the rapid and fine-grained exploration of parameters like scaling and rotation. For instance, to identify whether a scaling of 98.2% or 98.3% better explains the data, one simply computes an alignment error between a set of a few thousand correspondences. This error can then be minimized by a standard non-linear least-squares method. In contrast, using cross-correlation entails resampling the template image (e.g., the *Hinode*/SOT-SP level 2 scan)
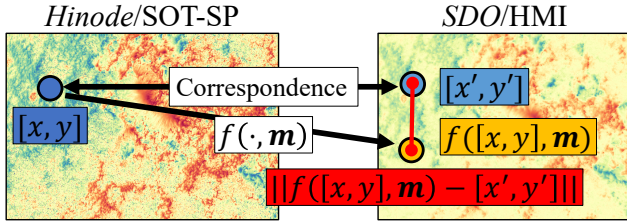
Figure 3. The matching setup. Given a correspondence $[x, y] \leftrightarrow [x', y']$, we aim to find a transformation f, parameterized by $\mathbf{m}$, that maps $[x, y]$ to closely match $[x', y']$. Without outliers (i.e., correspondences that are drawn at random), one can safely minimize the squared distance $||[x', y'] - f([x, y], \mathbf{m})||_2^2$. With outliers, one instead maximizes the number of correspondences satisfying $||[x', y'] - f([x, y], \mathbf{m})||_2 < \epsilon$.

and performing a cross-correlation. As an additional benefit, one can integrate multiple potential sources of information without having to account for differences in the scaling of the quantities; instead, one just concatenates correspondences from each data source.

Our full assumed transformation takes the form of a scaling, rotation, and translation, or

$$f([x, y]) = \mathbf{R}(\theta, c_x, c_y) \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}, \quad (1)$$

where $\mathbf{R}(\theta, c_x, c_y)$ is a rotation by θ about a point c_x, c_y (assumed to be the projection of 0°,0°) Corresponding points $[x, y] \leftrightarrow [x', y']$ are given, and the remaining parameters are either fixed to presumed values (e.g., s_x can be obtained by the relative scaling of the two sensors) or fit to data to minimize the squared Euclidean norm of their difference, or $||f([x, y]) - [x', y']||_2^2$. In particular, given a set of correspondences and any fixed parameters, one can solve for the free parameters that best ensure equation (1) matches each $[x', y']$ in the least-squares sense via nonlinear optimization to minimize the sum of squared distances between the correspondence location (i.e., $[x', y']$ and modeled correspondence location (i.e., $\mathbf{R} \operatorname{diag}([s_x, s_y])[x, y] + [t_x, t_y]$). Throughout, we fit transformations using Levenberg-Marquardt (Levenberg 1944; Marquardt 1963).

The model expressed in equation (1) fits five parameters: rotation angle θ, anisotropic scaling s_x, s_y, and translation t_x, t_y. In theory, the scaling parameters s_x and s_y are fixed parameters that do not change from scan to scan. The other parameters, θ, t_x and t_y are fit per-scan and indicate the relative pointing information. However, by fitting s_x and s_y, we can investigate what scale parameters best explain correspondence data. We additionally analyze how well models that assume different s_x and s_y explain the correspondence data. Finally,

with s_x and s_y fixed, one can use the θ, t_x and t_y fit to correspondences to estimate pointing information.

For convenience, we also use a simplified transformation consisting of an assumed lack of rotation (i.e., $\theta = 0$), an *assumed* scaling, and a fit translation. In particular, this transformation assumes that

$$f([x, y]) = \mathbf{R}(0, c_x, c_y) \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}, \quad (2)$$

where s_x, s_y are fixed parameters and $\mathbf{R}(0, c_x, c_y) = I$. The only free parameters are a translation vector $[t_x, t_y]$, which can be solved for in closed form with a single correspondence $[x, y] \leftrightarrow [x', y']$. The transformation of equation (2) is useful for the fast screening of outliers: its assumption of $\theta = 0$ is fairly close to reality and so it usually describes correct correspondences fairly accurately. Since it can be fit to a single correspondence, one can simply enumerate all correspondences to examine potential translations supported by the data.

For validation, one can also fit an affine transformation, which takes the form

$$f(x, y) = \mathbf{A} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (3)$$

where the parameters $\mathbf{A} \in \mathbb{R}^{2 \times 2}$ with no constraints and $[t_x, t_y]$ are fit to data. The transformation is slightly incorrect since it has an extra degree of freedom: due to the SVD, $\mathbf{A}$ can be always factored as $\mathbf{U} \operatorname{diag}([\sigma_1, \sigma_2]) \mathbf{V}^T$ where $\mathbf{U}, \mathbf{V}$ are rotations, and accordingly this transformation is equivalent to a rotation, scaling, and rotation followed by a translation. This extra degree of freedom enables the affine transform to also include a shearing. While the transformation is incorrect, it is useful since it can provide validation of the fits that is independent of any errors with regards to the center of rotation c_x, c_y. Moreover, the fit can be solved in closed form since equation (3) can be fit directly using ordinary least-squares.

3.4. *Robustly Fitting Models*

The correspondences found in Section 3.1 are contaminated with large numbers of outliers, which poses a problem for model fitting. In particular the transformation models described in Section 3.3 are fit using least-squares and are therefore highly sensitive to outliers. We overcome the challenge of outliers via RANSAC (Fischler & Bolles 1981), a simple but effective mechanism for model-fitting in the presence of the outliers.

Suppose we are given a set of N correspondences $[x_i, y_i] \leftrightarrow [x'_i, y'_i]$, a model with parameters $\mathbf{m}$, and a forward modeling function f such that $f([x_i, y_i], \mathbf{m})$ should

ideally match $[x'_i, y'_i]$. For example for the affine model (Eqn. 3), the parameter vector $\mathbf{m}$ is an element of $\mathbb{R}^6$ and encapsulates both $\mathbf{A}$ and $[t_x, t_y]$. Then $f([x_i, y_i], \mathbf{m})$ computes $\mathbf{A}[x, y] + [t_x, t_y]$. Least squares aims to find a model that minimize the sum of squared errors, i.e., $\sum_{i=1}^{N} ||[x'_i, y'_i] - f([x_i, y_i], \mathbf{m})||_2^2$.

Instead of trying to explain *every* data point well, RANSAC aims to maximize the number of correspondences that are well-explained by the model, or $\sum_{i=1}^{N} ||[x'_i, y'_i] - f([x_i, y_i], \mathbf{m})||_2 \leq \epsilon$ for some empirically set parameter ϵ. This set of points that are well-explained by the model is referred to as the model's *inliers* (in contrast to outliers). RANSAC proceeds by generating and then checking k hypothesized models $\mathbf{m}_1, \ldots, \mathbf{m}_k$. For each hypothesized model $\mathbf{m}_j$, one computes the inliers associated with $\mathbf{m}_j$ (i.e., $\{i : ||[x'_i, y'_i] - f([x_i, y_i], \mathbf{m}_j)||_2^2 \leq \epsilon\}$). The model with the most inliers is considered the "best model", and the best model's inliers are assumed to be inliers for the true transformation. Therefore, one usually re-fits the model in a least-squares sense on the best model's inliers, or $\arg\min_{\mathbf{m}} \sum_{i \in \mathcal{I}} ||[x'_i, y'_i] - f([x_i, y_i], \mathbf{m})||_2^2$ if $\mathcal{I}$ are the indices of the inliers.

The hypothetical models $m_1, \ldots, m_k$ are generated by fitting models on small sets of correspondences that are selected via exhaustive or random search. RANSAC fits hypothetical models on small numbers of correspondences, which minimizes the chance that an outlier is included and contaminates the fit. For some models, one can directly enumerate the space of potential models. For instance, the simplified model of equation (2) has two free parameters (for translation) and can be exactly fit on one correspondence. Thus, one can simply generate a set of hypothetical models by trying the translation fit on each of the N correspondences. In other cases, (e.g., the affine model in equation 3), the space of possible models is too large: since the affine model has six parameters and each correspondence provides two equations, model fitting requires at least 3 correspondences. With large numbers of correspondences, $\binom{N}{3}$ is too large to search directly, and so one resorts to randomly sampling. Specifically, one samples k triplets of correspondences without replacement, where k is an empirical parameter.

Practically speaking, this means that a RANSAC fit of an affine model consists of: (1) randomly sampling k triplets of points without replacement and fitting an exact model m_j and counting how many correspondences are well-described by m_j; (2) taking the inliers of the model with the most inliers, and then performing a least-squares fit of equation (3) on these inliers.

3.5. *Accounting for Non-Uniform Slit Motion*

As described in Section 2, the $H \times W$ Level 2 magnetogram produced by *Hinode*/SOT-SP is shorthand for a larger $H \times W'$ magnetogram ($W' \geq W$) with missing columns representing slit positions that have been skipped. The spectrograph slit position defines a mapping from each column in the original image $1, \ldots, W$ (representing a scan index) to another column in the larger image $1, \ldots, W'$ (representing a slit position). We use this mapping to interpolate the column indices of the locations in the scan indexed image, enabling a fit to be performed on the data as if the x coordinate referred to a slit position. Thus, a feature for correspondence found at x location pixel 45 may be relocated to pixel 49 if there were four missing columns at the start of the scan. The methods used (Lowe 2004; Rublee et al. 2011) estimate the locations of regions of interest at sub-pixel locations, which are handled similarly: a location at pixel 45.7 would be relocated to 49.7.

The use of the correspondences where x refers to the slit position rather than the scan number requires only two small changes when one transforms data between *Hinode*/SOT-SP level 2 data and *SDO*/HMI. When one warps *SDO*/HMI to the *Hinode*/SOT-SP scan index grid, one first must warp to the grid defined by slit positions, and then drop the missing columns. When one warps *Hinode*/SOT-SP data to the *SDO*/HMI grid, one must first fill in the missing columns. Our later analysis excludes scans with substantial numbers of missing columns, so this linear interpolation is unlikely to cause substantial difficulties.

3.6. *The Final Models and Model-Fits*

These fitting procedures begin with inlier identification via RANSAC, which produces a best model and corresponding best inliers. Inlier identification is followed by least-squares fitting, in which a model is fit to the best inliers via least-squares.

1. The full model fits $\theta, s_x, s_y, t_x, t_y$ to the data after identifying inliers with the simplified model, assuming the theoretical Level 2 values of s_x, s_y. To find the simplified model, one fits and then tries all hypothetical models; these models are generated by all the correspondences. This simplified model's inliers are then refitted with the full model (i.e., $\theta, s_x, s_y, t_x, t_y$) by nonlinear least-squares; the full model is initialized with the translation vector from the best-fitting model, the theoretical Level 2 values for s_x, s_y, and a rotation of 0.
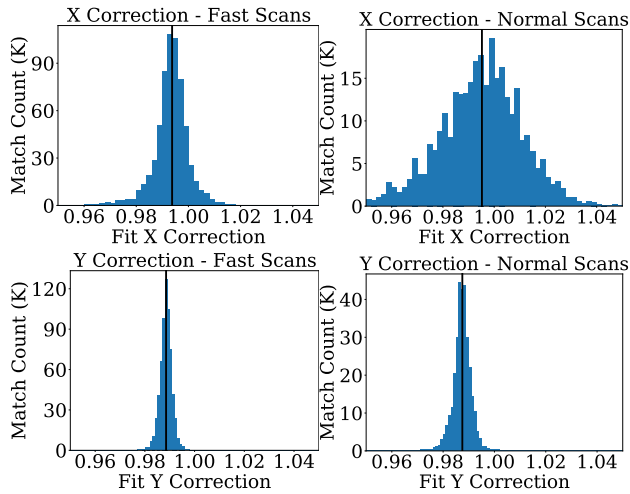
Figure 4. Histograms of the fit scale correction factor, weighted by the number of inlier matches. The 25% trimmed mean is drawn on the histogram as a black vertical line. There is a strong preference for X and Y scalings that are less than the theoretical Level 2 value on both normal and fast scans. There is substantially more scatter in the X direction, especially for normal scans.

2. The theoretical Level 2 model fits θ, t_x, t_y and with s_x, s_y fixed at their theoretical values (i.e., the relative scaling of `XSCALE` using the values in the headers). Like the full model, one identifies a set of inliers via the simplified transformation while assuming $\theta = 0$. These inliers are used for a non-linear least-squares fit for θ, t_x, t_y that is initialized with $\theta = 0$ and the translation set to the best model from RANSAC.

3. The empirical model is identical to the theoretical model, except it assumes a given empirical correction to s_x and s_y that we identify in Section 4.

4. The affine model that fits $\mathbf{A}, t_x, t_y$. We identify a set of inliers by fitting $100K$ hypothetical affine models on random triplets from the dataset. We then fit an affine model via least-squares on the inliers of the best model. After $\mathbf{A}, t_x, t_y$ are fit to data, $\mathbf{A}$ can subsequently be factored via the SVD to give two scales.

Throughout, our definition of an inlier is a correspondence where the error $||f([x, y] - [x', y']||$ is less than $\epsilon = 1$ HMI pixel (i.e., $\approx 0.5''$). When fitting a model with no rotation for pre-screening inliers, we use a threshold of $\epsilon = 2.5$ HMI pixels (i.e., $\approx 1.25''$) since this model is inaccurate due to its failure to account for rotation.

4. ESTIMATING RELATIVE SCALE

Table 1. Fit scales (in arcseconds) compared with Level 1 and Level 2 reported scales. The Level 1 y scale is from `CDELT2` and the x scale is the nominal stepping size scaled appropriately (1x for normal scans and 2x for fast scans). The fit scales match the Level 1 data better than the Level 2 data.

	X Scale			Y Scale		
Scan	Fit	L1	L2	Fit	L1	L2
Fast	0.2953	0.2952	0.2971	0.3163	0.3170	0.3199
Normal	0.1479	0.1476	0.1486	0.1580	0.1585	0.1599

Table 2. Fit Scales. In the top two rows, we report percent of the nominal Level 2 value. In the bottom two rows, we report the updated scales in arcseconds. We report the 25%-trimmed mean of the scales fit via the full model, the affine model (that does not assume a particular rotation center), and their difference. The affine model results in a similar conclusion, that the scale is smaller than is reported.

	X Scale			Y Scale		
	Full	Aff.	Diff.	Full	Aff.	Diff.
Fast	99.37	98.95	0.43	98.85	98.45	0.41
Normal	99.52	99.21	0.32	98.75	98.50	0.25
Fast	0.2953	0.2940	0.0013	0.3163	0.3150	0.0013
Normal	0.1479	0.1474	0.0005	0.1580	0.1577	0.0004

We begin by identifying the relative scale of the two instruments' observations. We first treat the scales s_x and s_y as a parameter to fit over the entire dataset, and report an estimate of their values as described by the correspondences. We then analyze whether the updated scales can better explain the data compared to the theoretical scale. Having fit the data, we examine whether there are spatiotemporal trends in the scales that were fit. We additionally consider and exclude some alternate explanations for our results. Throughout this analysis, we use the Fast and Normal subsets of the data; the full dataset is used for pointing information.

4.1. *Estimating Fitting*

We first fit the full model (which has scaling as a free parameter) on the *Fast Subset* and *Normal Subset* dataset. Some fits fail, and therefore we report the 25-% trimmed mean (i.e., the mean of the middle 50% of the data). We found the results to be insensitive to whether

we took the mean scale, weighted by correspondence or not, and to various trim-levels. We plot histograms of the fit scales in Figure 4, weighted by the number of matches, and plot the 25%-trimmed mean. In the Y-direction, both fast and normal scans show a substantial preference for a scale at near 98.8% the theoretical Level 2 value, with little scatter. In the X direction, there is substantially more uncertainty, but a preference for scale near 99.4% of the theoretical Level 2 value. We report the fit scales compared to the Level 1 and Level 2 reported values in Table 1. The data show substantially better agreement with the data from *Hinode*/SOT-SP Level 1: fitting finds a value that is 100.2% of the nominal Level 1 value in Y and 99.7% of the nominal Level 1 value in X.

To exclude the possibility that the full model is poorly specified e.g., that something is wrong with the rotation location – we analyze the inferred scales of the affine model. We factor the fit affine models into a rotation, scaling, and rotation via the SVD, and report the fit scales in Table 2. The results are virtually identical, with relative changes in scale of at most 0.41%. Moreover, the affine model consistently estimates a *smaller* scale factor.

To help quantify the adjustment of scale factors, we convert it to pixels. An adjustment of 1.2% to the Y scale factor of amounts to ≈3.8px error over a typical 320 *SDO*/HMI pixels that a *Hinode*/SOT-SP scan covers when warped to the *SDO*/HMI grid. However, with a model that is fit to describe a mapping from a *Hinode*/SOT-SP scan to a smaller cutout on *SDO*/HMI, some of the error can be compensated for by adjusting the translation of the center of the window. This leads to ≈1.9px error over a typical *SDO*/HMI window. This calculation suggest that *SDO*/HMI pixel-sized errors can cause a scale adjustment of ≈0.625%.

4.2. *How well do the data fit?*

We then turn to testing how well models explain the data when they assume different relative scales. While the data are fit to correspondences, once the transformations models have been fit, they can be used to warp the full *Hinode*/SOT-SP scan to the *SDO*/HMI grid. The warped *Hinode*/SOT-SP data can then be compared with the corresponding *SDO*/HMI grid, and its agreement can be used as a measurement of goodness of fit. In particular, we compare a model that assumes the theoretical scales given by the Level 2 header keywords (the *Theoretical* model) and one that assumes the empirically fit scale with the corrections found in this paper (the *Empirical* model).

Table 3. Comparison of the Spearman ρ rank correlation between *SDO*/**HMI** data and *Hinode*/**SOT-SP** data, warped with either an Empirical model, or the Theoretical Level 2 model. Results presented average over field, inclination, and continuum, but trends hold for each individual quantity. Ties are defined as within the 5% of the the average rank correlation.

Scan Type	Empirical Better	Tied	Theoretical Better
Fast	58.7	41.3	0.0
Normal	27.6	72.4	0.0

Table 4. Comparison of the extent of the features that are well-explained for the empirically fit and theoretical Level 2 parameters. We report the fraction of times the empirical model covers more of each axis by at least 10 pixel (*More*); the fraction that the extents are within 10 pixels (*Tie*); and the fraction of times the theoretical model explains a larger fraction of the image (*Loss*). The empirical correction often explains correspondences across more of the image, and rarely less. Calculated on matched scans where both have at least 20 inliers.

	X			Y		
Scan	More	Tie	Less	More	Tie	Less
Fast	56.5	38.6	4.9	88.3	10.6	1.2
Normal	17.3	73.8	8.9	73.4	24.5	2.0

Throughout, we produce the *Hinode*/SOT-SP data warped onto the *SDO*/HMI grid by interpolating with a degree 3 bivariate spline. As previously described, most *Hinode*/SOT-SP scans have non-uniform slit motion and therefore have missing scanlines. We interpolate these values by linearly interpolating between the recovered scanlines; while a poor approximation over large discontinuities, the data deliberately only includes slits with minor discontinuities.

We evaluate the rank correlation between *SDO*/HMI data and the warped *Hinode*/SOT-SP data, comparing continuum, inclination, and field strength (*SDO*/HMI: `field`; *Hinode*/SOT-SP: the product of `Field_Strength` and `Stray_Light_Fill_Factor`). In this case, since we are doing pixel-to-pixel comparisons, we explicitly compute magnetic flux density for *Hinode*/SOT-SP. Using rank correlation as a measure of agreement compared to e.g., distance, avoids issues of relative calibration and scaling of the two pipeline quantities. While the rank correlation itself is not infor-
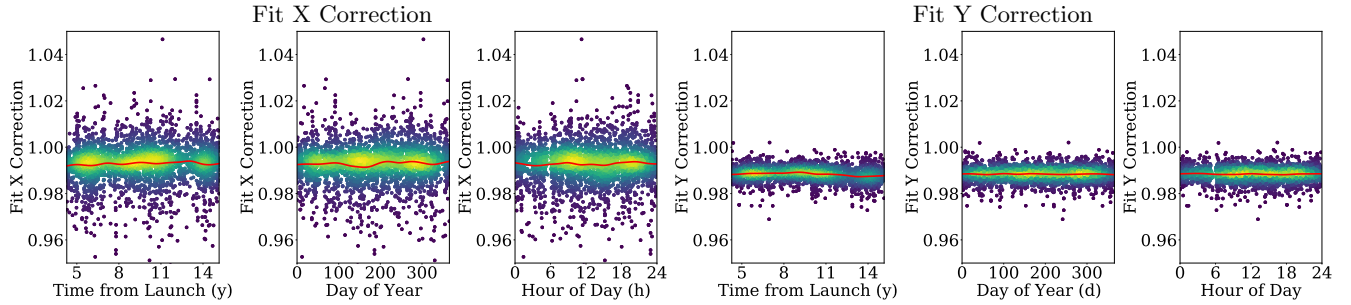
Figure 5. Scatterplots of fit X and Y corrections against time since launch (22 September 2006), day of year and time of day. The color of each scatterplot point indicates the density at that location, estimated with a kernel density estimate. We additionally plot a kernel regression estimate in red of scale correction as a function of each covariate. The corrections have no trend, suggesting that the scale factor is constant and that causes driven by Earth or satellite orbit are not responsible.

mative as an absolute quantity, the relative rank correlations can be used to sort different methods: one can check how many times one transformation yields data with a higher rank correlation compared to another.

We use this approach to compare the fits produced with the theoretical Level 2 model and the empirical model. We average over the three quantities and for each *Hinode*/SOT-SP scan, we pick at random from the HMI transformations for which the model has at least 20 inliers. Table 3 shows that the empirical correction results in warped data that is never worse than the theoretical values, and often better.

Another method of testing goodness of the fit scale is quantifying the extent to which inliers for each model cover the image. We quantify performance by examining the spatial extent of the inliers of the theoretical as well as the empirical model. A mis-specified scaling can be corrected for in a small region by adjusting the translation. However, changes in translation cannot compensate over a large region.

We quantify the extent via the size of the middle 95% of the data for a set of inliers $[x_i, y_i] \leftrightarrow [x_i', y_i']$. For x, this is quantile($\{x_i\}, 97.5\%$) − quantile($\{x_i\}, 2.5\%$) and similarly for y. We compare the rate at which the empirical model's inliers cover a wider fraction of the image compared to the theoretical model's inliers and report results in Table 4. The empirical model usually covers more of the image, suggesting that the empirical model is a better description for more of the image compared to the theoretical Level 2 values. The difference is consistently visible in the Y direction, and larger for fast scans compared to normal scans. We hypothesize that this difference is caused by a difference in the field of view covered by the scans that could be registered: we found that the fast scans that could be registered typically had a larger field of view compared to the normal scans that could be registered. Since the error due to an incorrect scaling increases linearly as a function of
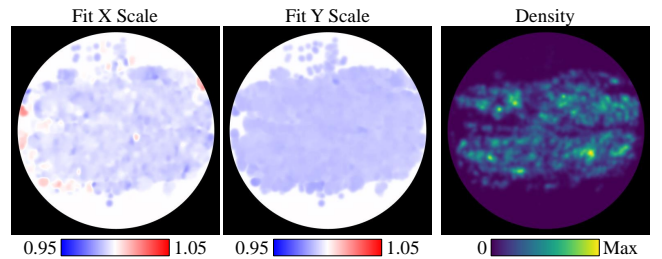

Figure 6. Fit scales by location. We plot a kernel regression estimate of the location, with a prior of 1. We also plot the density. The apparent bias in density towards one side of the disk is driven by the *Hinode*/SOT-SP observing cycle: an active region needs to have become previously visible in order for *Hinode*/SOT-SP to have it selected as a target. The corrections also have no trend (although there are some outliers), suggesting that the location-dependent causes such as differences in line formation height are not responsible for the change in scale.

field of view, the fast scans likely have more room for improvement with better alignment.

4.3. *Identifying Correlated Signals*

We next turn to identifying whether the estimated scale correlates with temporal or spatial quantities.

For temporal trends, we examine both a secular trend in scale over a decade (similar to the instrument degradation in e..g, *SDO*/AIA's data) as well as cyclic trends in the data. In particular, we test yearly cycles (to account for artifacts related to the Earth's orbit) and daily cycles (to account for artifacts related to orbit such as those in *SDO*/HMI's data (Hoeksema et al. 2014). We plot the fit scale corrections against these variables in Figure 5 along with a kernel regression estimate of the fit scale as a function of the covariate. Given a set of N locations $\{\mathbf{l}_i\}_{i=1}^N$ with values $\{v_i\}_{i=1}^N$, and a kernel function k, the value of the kernel regression at a query point $\mathbf{q}$ is the values weighted by the kernel distance to the query point, or $(\sum_{i=1}^N k(\mathbf{q}, \mathbf{l}_i) s_i)/(\sum_{i=1}^N k(\mathbf{q}, \mathbf{l}_i))$. This
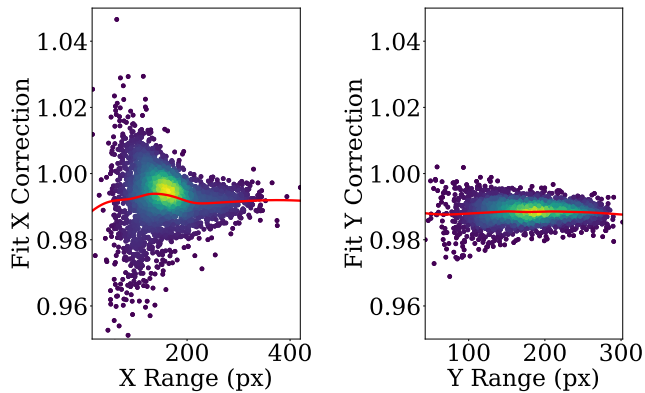
Figure 7. Scatterplots of fit X and Y corrections against the range of the X and Y coordinates. Having more correspondences across the full image results in more stable estimate of the scales in the X coordinate.

Table 5. Fit Scales for both all our data and near-smooth scan data. We report scales (in %) on both all the scans from our fast dataset that can be registered with at least 20 correspondences explained within 1px, as well as the subset of these where the number of excess slit positions is no more than two positions. We report the 25%-trimmed mean of the scales fit via the full model. We reach the same conclusions when using the smaller number of scans with close-to-no jumps.

	All Scans			≤ 2 Extra Slit Positions		
Scan	X	Y	Count	X	Y	Count
Fast	99.37	98.85	760	99.38	98.82	287
Normal	99.52	98.75	503	99.95	98.70	98

kernel regression smooths out the noise in each individual fitting. We find no trends in any of the temporal data. This lack of trends would suggest that the pixel scale has not changed over time, which agrees with the intuition that pixel scale is an unchangeable property (at least at the precision we can measure).

Next, we examine the estimated scale as a function of image plane location in Figure 5 to see if there are spatial trends across the disk (e.g., a hemispheric bias or correlation with distance-from-center). We compute a kernel regression estimate, using the HMI coordinates as locations and the fit scales as values. We use a Gaussian kernel with $\sigma = 64$px, or $k(\mathbf{x}, \mathbf{y}) = \exp(-||\mathbf{x} - \mathbf{y}||_2^2/\sigma^2)$. Many parts of the disk are far from any data point, so for regularization, we also add a single dummy correspondence at the query location with scale 1, leading to a scale of 1 where there is no data. The data are relatively consistent across the disk. While some spots on the disk deviate from the overall average, these are likely due to a handful of scans, rather than systematic biases.

Finally, we plot the fit corrections against the range of the inlier correspondences in x and y in Figure 7. We define the range via the range of the middle 95% of the data. Unsurprisingly, as feature correspondences cover less of the image, the correction that is fit exhibits substantially more variability.

4.4. Analysis of the Correspondences

For fast data, the bulk of the inlier matches come from Inclination (39.1%) and Field (36.6%), with continuum (17.2%), and V channels (7.0%) making up most of the rest. Most (83.7%) of the correspondences come from SIFT, with the rest from ORB (although we note that the SIFT detector configuration we used produced more

correspondences to start with). Among the image pairs where there are at least 20 inlier correspondences, on average 6% of the data are inliers and 94% outliers. This fraction is driven by both the challenging nature of the corresponce problem as well as the fact that a relatively small cutout from *Hinode*/SOT-SP is being matched to a much larger full disk image from *SDO*/HMI.

The high outlier fraction (94%) seems daunting at first. However, two things facilitate the extraction of good models. First, the outliers are random, as opposed to structured. This means that RANSAC (Fischler & Bolles 1981) cuts through them with relative ease. Second, although our models that fit rotation and scale require 3 correspondences to be fit, we use a fast pre-screening model that fits to a *single* correspondence while using an assumed rotation and scale. This is only possible because $\theta = 0$ is a good approximation for reality and the reported scale parameters are already quite accurate. This insight has also been applied in the context of field robotics by Civera et al. (2010), where vehicle motion is used to constrain the search for more complex models.

4.5. Excluded Explanations

We now briefly discuss factors that we exclude as potential causes of the observed scale factor. These alternate explanations include the scale factor being used to explain away temporal evolution, the relative locations of both the instruments and the phenomena they observe, and issues with the accounting for the mechanical slit position. We cannot rule out that these factors *cause* changes in scale factor because each obviously results in some relative scaling. Instead, we want to demonstrate that each is too small to plausibly explain the observed changes and would have left a signature in our results.
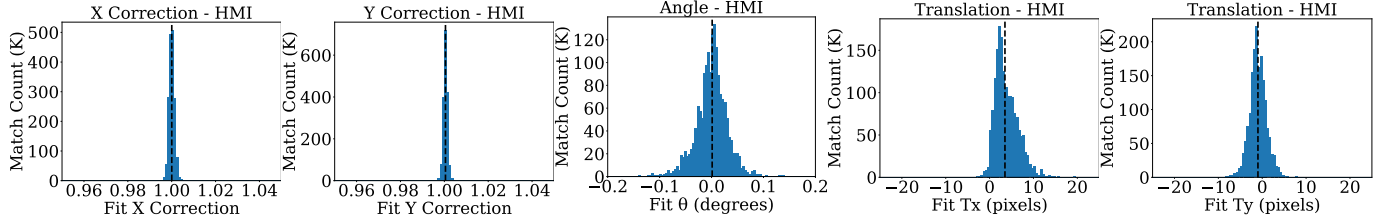
Figure 8. Fit on temporally adjacent *SDO*/HMI scans from the `ME_720_fd10` series. In both cases, the fit scales are effectively unity and rotation zero. Translations have substantial scatter with the x axis centered on a positive value, and the y axis centered on zero. This agrees with intuition, suggesting that the fitting procedure is effective even with temporal evolution.

One hypothesis for the observed fits is that solar evolution during the *Hinode*/SOT-SP scan is responsible for the change in scale. If one rotates the coordinate system by $\theta \gg 0$, a change in the y scale can compensate for changes in the x direction. We therefore test the fitting procedure by registering temporally adjacent HMI scans, and observing the fit parameters and rotations.

The relationship between the pixels in temporally HMI scans *across the full image* is not described by any of the models used in this paper. However, locally within an active region, these transformations are a reasonable model. We therefore perform fits on the same active regions imaged by *Hinode*/SOT-SP, which we identify by mapping the *Hinode*/SOT-SP scan warped to the *SDO*/HMI grid and computing its bounding rectangle. We find correspondences between *SDO*/HMI's field, azimuth, inclination, and continuum, filtering the correspondences to lie in this rectangle. There is no scaling or rotation between temporally adjacent *SDO*/HMI scans, and so we should expect s_x and s_y to be near unity and θ to be near zero. The Sun's rotation ought to produce usually positive translation in the x component. Because the Sun is rotating out of the image plane, the y rotation will not be zero in each image; however, unless there are biases in data sampling, it should have mean zero over the dataset. We report results in Figure 8. The fit scale corrections lie close to unity. The Y scale has a 25% trimmed mean of 100.05%, and 95% of the data falls between 99.86% and 100.24%. The X scale has a 25% trimmed mean of 100.00% with 95% of the data between 99.77% and 100.25%. The fit rotation tends to have a magnitude near zero. Translation has substantial scatter, but Y is centered on zero and X is positive.

Another possible explanation is the relative positioning of the spacecraft carrying the instruments. However, they are insufficiently far apart to explain the change in scale we observe: SDO's orbit is less than $\approx 3.6 \times 10^4$ km and Hinode's orbit is less than 7×10^2 km. The sum of their orbits gives an upper bound on their potential distance, yielding $\approx 3.7 \times 10^4$ m, or just 0.025% of the average Earth-Sun distance (1.5×10^8 km). This fraction of the distance is negligible compared to the scale

diferences seen ($\approx 1.18\%$ and $\approx 0.64\%$). Additionally, if spacecraft position were a factor, we would expect to see variation in scale as a function of time of day via the interaction between both spacecrafts' orbits or time of year. The plots in Figure 5 show no such trends.

In addition to being at different locations, the instruments observe different parts of the photosphere. These height differences are also insufficiently different to explain the scale differences. Higgins et al. (2022) report a theoretical estimate of no more than $\approx 1.5 \times 10^2$ km difference of formation heights using results given in (Grec et al. 2010; Norton et al. 2006). This height difference results in at most a sub-pixel misalignment. This upper bound can be computed by making assumptions that maximize misalignment. We compute the 150km height difference at the limb and assume the Sun takes up the full sensor, yielding a radius of 2048 HMI pixels. In this setup, each pixel amounts to ≈ 339km per HMI pixel. The error due to line height is thus ≤ 0.443px, which is a loose upper bound given that line formation height difference has no impact at disk center. This error, however is substantially smaller than the translation error induced by a scale change of 1.9px. Moreover, the difference in pixel location due to formation height varies as a function of distance to disk center. If formation height were the cause of scale differences, it would leave a signature of substantial variation of the fit scale factor across the disk. However, the plots in Figure 6 show a near constant function for Y and while X is noisier, no consistent trends emerge.

We finally consider jumps in the scanning slit mechanism. While we account for these in our fitting, we re-run the analysis on a subset of scans where there are nearly no jumps. In particular, we consider only scans with two or fewer extra slit positions, corresponding to $0.297''$. For *SDO*/HMI, this amounts to ≈ 0.5 pixels. For *Hinode*/SOT-SP, it is one pixel in fast mode, and two pixels for normal mode. We report results in Table 5. Our results show good agreement with the full dataset with the exception of the scale fit for Normal scans in the X direction. This quantity is, however, both noisy and fit on a small number of scans.
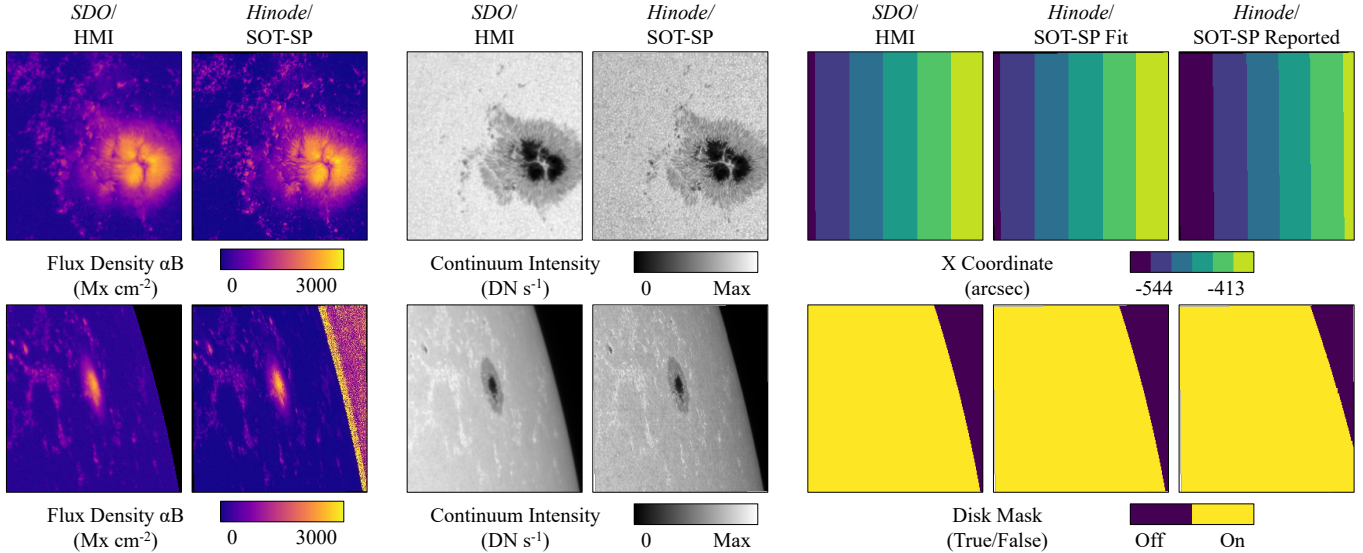
Figure 9. Warped data and pointing coordinates. We show a subset of the maps for conciseness, including flux density (αB) and continuum intensity and coordinate maps and disk masks. Coordinates shown for *fit* values that are produced by the fit alignment with *SDO*/HMI and *reported* values that are reported in the Level 2 *Hinode*/SOT-SP data. The reported values show substantial disagreement with *SDO*/HMI data even though the alignment of continuum and field strength are close to pixel-perfect. This is most clear by examining the limb pixels on bottom. In comparison, our fit corrections align well. Top: 2011 April 30, 09:55:06 TAI. Bottom: 2013 April 3, 15:47:05 TAI.

5. ESTIMATING RELATIVE POINTING

Hinode/SOT-SP pointing information is known to be inaccurate, as described in its data analysis guide (Hurlburt et al. 2009, Appendix C). However, the transformations fit between *SDO*/HMI and *Hinode*/SOT-SP let us provide updated relative pointing information. This pointing information is relative to *SDO*/HMI, and is therefore limited by the accuracy of *SDO*/HMI's pointing information, any optical distortions (Schou et al. 2012), and the accuracy of our procedure. However, *SDO*/HMI's observations include the limb of the Sun, which permit a more accurate determination of its pointing. We now explore updating *Hinode*/SOT-SP's pointing information using the full 2011-2021 dataset.

Just as one can warp *Hinode*/SOT-SP data to the *SDO*/HMI grid to enable applications such as the synthetic inversion system presented in (Higgins et al. 2022), one can also warp a grid of coordinates. The reported coordinates are provided in the Level 2 *Hinode*/SOT-SP data products as `X_Coordinate` and `Y_Coordinate` as well as implicitly via the values `XCEN` and `YCEN`. To update the coordinates, we warp the helioprojective coordinates from *SDO*/HMI, which we are calculated via SunPy (The SunPy Community et al. 2020). The transformation that was fit between *SDO*/HMI and *Hinode*/SOT-SP references the slit position as opposed to the scan index; however, recall that this can be transformed to the Level 2 *Hinode*/SOT-SP data by dropping slit positions that were not observed.

Our procedure does not account for relative spacecraft position, but we estimate that the relative position leads to an error of under $1''$.

We illustrate this warping in Figure 9, showing alignment for flux density (αB), continuum intensity, and disk masks. The *Hinode*/SOT-SP data is warped to the *SDO*/HMI grid according to the fit transformation using a cubic spline for interpolation. The *Hinode*/SOT-SP field strength is calculated by the product of intrinsic field strength `Field_Strength` and fill factor `Stray_Light_Fill_Factor`. A transformation that aligns both instruments' field strength (or similarly, their continuum images) results in substantial misalignment of the reported *Hinode*/SOT-SP information but good alignment with our corrections (by construction).

Our updated coordinates can then be used to update pointing information `XCEN` and `YCEN`. We note, however, that a few scans have substantially incorrect data for `XCEN` and `YCEN` in the header but more accurate x and y coordinates in the `X_COORDINATE` and `Y_COORDINATE` fields; we thus recalculate `XCEN` and `YCEN` by the average of the x coordinates in the first and last columns, and similarly for the rows and the y coordinate. We calculate the updated values of `XCEN` and `YCEN` identically using our updated coordinates that are warped from *SDO*/HMI. This usually does not change coordinates, except in the small number of cases with substantial mistakes.
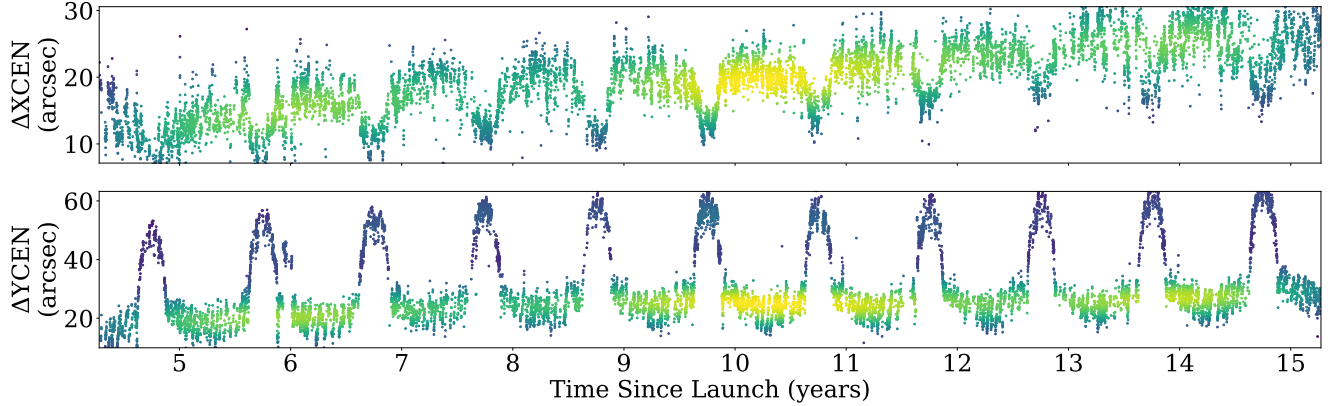
Figure 10. Pointing updates as a function of time since launch (22 September 2006). We plot ΔXCEN and ΔYCEN as scatterplots (showing the middle 98% of the data), coloring points by their density using a kernel density estimate. There are substantial secular and cyclic trends. This plots uses improved estimates of XCEN and YCEN calculated via the X_COORDINATE and Y_COORDINATE fields to update some incorrect (but easily fixed) values reported in headers. Not doing this correction would mix a large easily-fixed error involving a small amount of the data with the few-dozen arc-second more difficult-to-remove trend impacting most of the data. The effects of these substantially incorrect XCEN and YCEN header values are shown in Figure 16. Density colormap: Min ▓▓▓ Max
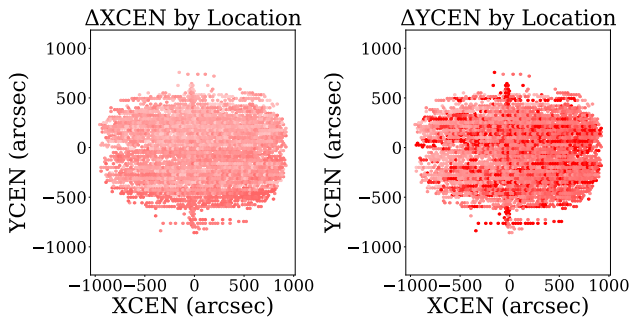
Figure 11. Plots of ΔXCEN and ΔYCEN as a function of XCEN and YCEN locations. For each location, we plot the averaged ΔXCEN (left), ΔYCEN (right) per location. Colormap: -50″ ▮▮ 50″. There is a strong positive bias, as indicated by the uniformly red color.

We compute the residuals (i.e., ΔXCEN = XCEN$_{new}$ - XCEN$_{reported}$ and similarly for ΔYCEN). We plot the residuals as a function of mission time in Figure 10 and disk location in Figure 11 (displayed as a hexbin plot, averaging within hexagonal bins, due to the large number of data points). The plot of (ΔXCEN, ΔYCEN) as a function of time reveal strong annual oscillations in pointing on the order of a 30″ as well as a slow secular drift across over time. When plotted across the disk, these residuals are relatively consistent. Across the the dataset, the mean norm of the residual is 20.4″ in x axis and 32.5″ in the y axis. Our updated X and Y coordinates are near perfectly aligned with *SDO*/HMI by construction (although they are dependent on the accuracy of our mapping).

5.1. *Exploring The Pointing Residuals*

Given the nature of the residuals shown in Figure 10, we then examine the pointing residuals ΔXCEN, ΔYCEN in more detail. We initially explored temporal trends such as total of time since launch (which captures secular trends) as seen in Figure 10 as well as time of year (which captures Earth-orbit effects) and time of day (which captures instrument orbit effects). We show ΔXCEN and ΔYCEN plotted against the covariates in Figure 12. Both residuals shift substantially during the middle of the year during the *Hinode* eclipse season. This shift is especially pronounced for the Y residual, but is also visible in the X as well. We note that this shift does *not* occur during *SDO*/HMI's biennial eclipse season. After spotting this relationship, we also plotted the residuals against temperature data reported at the camera electronics box (CEB) and CCD, which are given by T_SPCEB and T_SPCCD in the Level 1 headers. There is a strong relationship between the residuals and temperature, especially at the CEB.

Similar relationships were observed by Mariska (2016) when performing co-alignment between *Hinode*/EIS and *SDO*/AIA, who also pointed to thermal issues. Like our results, the range reported in ΔXCEN and ΔYCEN were around 30″and 60″ respectively as reported by The Hinode Review Team et al. (2019). *SDO*/AIA and *SDO*/HMI do not share optical pathways, nor do *Hinode*/EIS and *Hinode*/SOT-SP. Moreover, the two instruments onboard *Hinode* do not have the same pixel size. The close agreement in the size of the effect despite different optical pathways and pixel sizes suggests that the
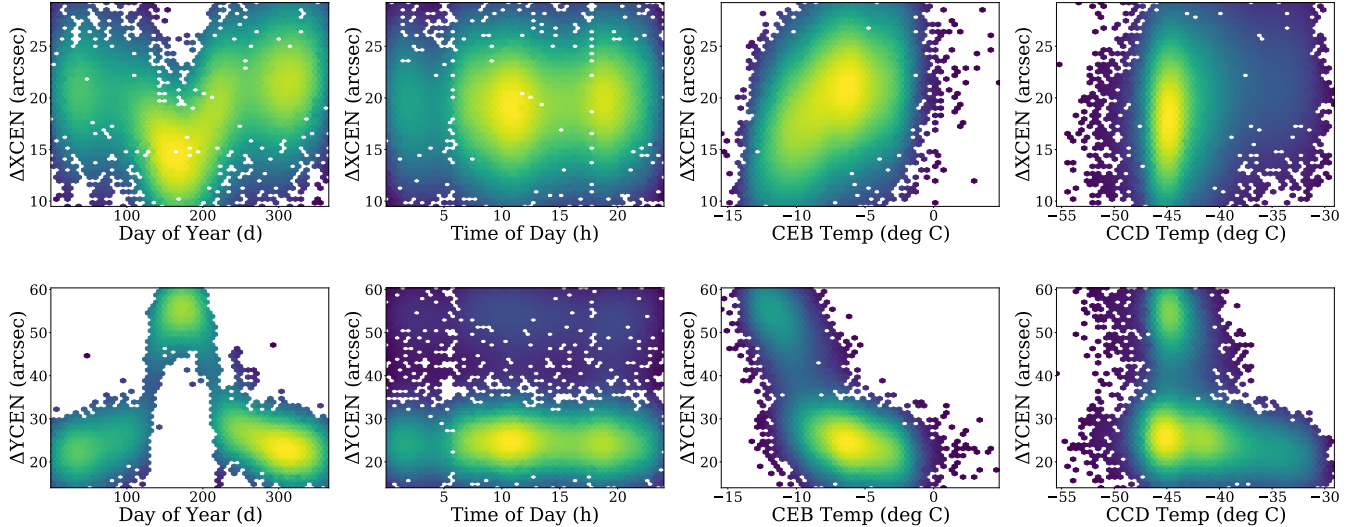
Figure 12. Plots of the fit ΔXCEN (top) and ΔYCEN (bottom) pointing residuals against multiple covariates. The colors of the points show the density in the area, estimated with a kernel density estimate. We plot the residuals against day of year and time of day, as well as temperatures at the Camera Electronics Box (CEB) and Charge-coupled device sensor (CCD). Both X and Y residuals have a large shift during eclipse season mid-way through the year. This shift is well explained by change in temperature, suggesting that thermal expansion plays a role. To avoid distorting the plot with clearly-fixable errors and a small number of clear outliers, the plots use XCEN and YCEN re-calculated from the coordinate fields, rather than the header and show the middle 98% of the data. Density colormap: White: Not observed; min ■■■■■ max.

thermally-driven variance in pointing originates in *Hinode*'s AOCS. We discuss this further in Section 7.

5.2. *Explaining The Pointing Residuals*

To more systematically explore the approach, we fit a linear model that takes covariates capturing time and temperature, and links them to the pointing residual (i.e., ΔXCEN and ΔYCEN). We use total time, time of year, time of day, and the CEB and CCD temperature. Each covariate was independently standardized (i.e., to have a mean of zero and standard deviation of unity). This prevents any regularization from being disproportionately applied to covariates with larger intrinsic ranges of values. The fit coefficients can explain the relative role of the parameters.

The relationship between many covariates and the pointing residuals is nonlinear, and so we performed a basis expansion for each covariate. This includes a quadratic term as well as well as radial basis function that that softly bins the covariate into a set of discrete categories across the typical range of values. Each radial basis function covariate is the probability of the value according to a set of Gaussians evenly spaced from -2 to 2. Thus, for a single covariate, the final result is the concatenation of the original covariate, its square, and the radial basis functions, or

$$\phi(x) = [x, x^2, p(x; \mu_1, \sigma), \ldots, p(x; \mu_{20}, \sigma)] \quad (4)$$

where μ_i is the ith Gaussian's mean, σ is an empirical parameter (set to 0.2), and $p(x; \mu, \sigma)$ is the Gaussian density.

Given the covariates $\{\mathbf{x}_i\}_{i=1}^{N}$ and a corresponding residual $\{y_i\}_{i=1}^{N}$ for either the x or the y axis, we solve for a linear model with an ℓ_1-regularized LASSO objective,

$$\underset{\mathbf{w} \in \mathbb{R}^F}{\arg\min} \lambda ||\mathbf{w}||_1 + \sum_{i=1}^{N} (y_i - \mathbf{w}^T \Phi(\mathbf{x}_i))^2 \quad (5)$$

where Φ applies ϕ to each coordinate of $\mathbf{x}$ and concatenates, producing an F-dimensional vector. The ℓ_1 penalty $\lambda ||\mathbf{w}||_1$ on the fit promotes simplicity in the form of sparsity and ensures that only a few covariates are selected. Rather than pick the trade-off parameter λ to maximize performance, we simply lower it until at most seven parameters are fit (picked empirically). We found that a few outliers (e.g., with pointing updates of $377''$) led to substantially worse fits due to the outliers dominating the fit. We therefore only fit on data points where *both* ΔXCEN and ΔYCEN were within the middle 99% of the data ($7.5'' \leq$ ΔXCEN $\leq 64.2''$ and $5.9'' \leq$ ΔYCEN $\leq 31.3''$). This removes of 206 out of the 12062 successfully registered scans.

The residuals are relatively well-explained by a subset of temporal and temperature-related variables. We report coefficients normalized so that the largest has value ± 1. The Y residual is explained primarily by radial
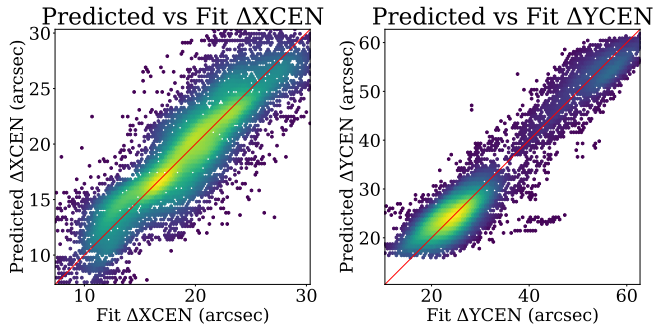
Figure 13. Pointing update prediction performance.
We plot the cross-validated predictions of ΔXCEN, ΔYCEN against the values fit by alignment. This plot shows the density of the region, estimated with a Kernel Density Estimate. The predicted values are generally quite accurate. Approximately 2% of data are outliers and are not shown. Density colormap: Min ▨ Max.

basis functions at mid-year, as well as a secular trend term. The fit model consists of: Time of Year radial basis function at day 180 (weight 1), another Time of Year radial basis function at day 158 (weight 0.78), CEB Temperature (weight -0.32), and Time Since Launch (weight 0.17), the squared CEB Temperature (weight 0.13) followed by squared covariates for Time of Year and Time since Launch, both with weights less than 0.05. The X residual is driven by a secular trend as well as a mid-year radial basis functions: the model is Time Since Launch (weight 1), a time of year radial basis function at day 158 (weight -0.83), then the CEB Temperature (weight 0.36), the square of the CEB Temperature (weight -0.11), with covariates for the time of year, squared time of day, and X scale getting weights of less than 0.05. We note that due to the strong correlation between CEB Temperature and time of year, the model may pick either. Once one variable is picked, the other one is not needed due to the ℓ_1 regularization.

5.3. *Predicting the Pointing Residuals*

Given importance of time of year and the total time and following Mariska (2016), we additionally fit a straightforward look-up table plus linear model consisting of two models. We refer to this as *Lookup-Plus-Linear*. The first model is a kernel regression (with $\sigma=3$ days) on the time of year, which functions like a smoothed look-up table to handle the bimodal behavior. The second model is an affine function (i.e., linear function plus bias) of the time since launch variable without any coefficient expansion. This linear function handles the secular trend found in the data. We fit the model stagewise: we make cross-validated kernel regression predictions to estimate the kernel regression mod-
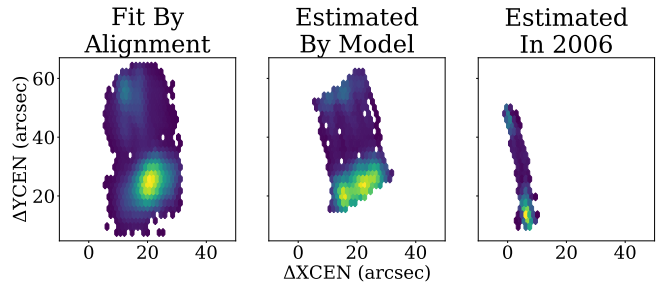

Figure 14. Pointing Residuals. (left) residuals fit via co-alignment; (middle) cross-validated predictions from temporal and thermal parameters; (right) cross-validated predictions on the same data, but with the data's timestamps shifted to approximately the time of launch. The predictions match the fit data fairly well but are pulled towards towards the mean (as is common minimizing the sum of squares). The extrapolations at time of launch shows a substantial reduction in residual. Since the predictions are cross-validated predictions (rather than the output of a single model), the time shift does not shift each datapoint identically. Approximately 2% of original fit residuals data are outliers (e.g., ΔXCEN = $300''$) and are not shown. Density colormap: Min ▨ Max.

Table 6. Mean absolute error (MAE) for prediction of the pointing residual. We report 20-fold cross-validated MAEs for four models for producing the pointing residual: (zero) using the data as-is; (median) using the median correction from the training set; (LASSO) an ℓ_1 model fit following §5.2; (Lookup+Linear) an additive model combining an interpolated lookup-table on time-of-year and a linear model on the time since launch. While the initial data is substantially misaligned, a simple model can reduce the misalignment by an order of magnitude.

Scan	Zero	Median	LASSO	Lookup+Linear
ΔXCEN	19.3	4.4	2.1	2.1
ΔYCEN	31.6	10.2	4.4	3.2

els' output and then fit a linear model on the residuals between the target and the cross-validated predictions.

The LASSO and lookup-plus-linear models explain the data well, as shown by the cross-validated MAE for multiple models shown in Table 6. While directly co-registering existing *Hinode*/SOT-SP scans to *SDO*/HMI data will lead to better alignment, these results suggests that one can apply a deterministic correction to substantially improve pointing information. The current pointing information is off on average by about $30''$ in Y. A constant correction reduces the error by 68%, and a simple model reduces the error by 90%. We show
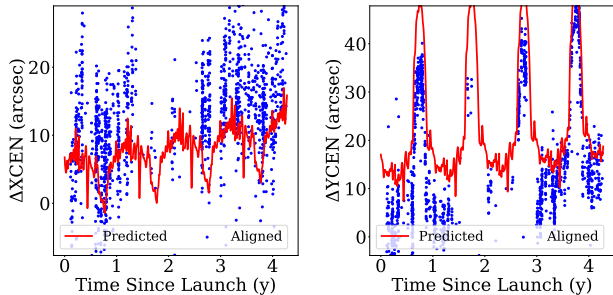
Figure 15.   Co-alignment of *Hinode*/SOT-SP with *SOHO*/MDI. We show a scatterplot of updated pointing information obtained from co-alignment with *SOHO*/MDI (in **blue dots**) alongside a prediction of the Lookup-Plus-Linear approach fit to the whole *SDO*/HMI dataset (as a **red line**). We plot both against time since launch (22 September 2006). As predicted by Figure 14, the pointing error via alignment gets substantially closer to zero near the launch date and continues to show the trends in Figure 10. The predicted pointing errors are reasonably accurate considering that this is an extrapolation.

cross-validated predictions against the fit residuals in Figure 13, which shows reasonably tight agreement.

While *SDO*/HMI was not available at the time of *Hinode*/SOT-SP's launch, we can use our fit models to extrapolate pointing residuals back in time to the launch of *Hinode* in 2006. We additionally predict the pointing residual on data points that have been altered by subtracting whole years from the time since launch, so as to leave other variables (e.g., time of year) unaltered. We plot these results in Figure 14, showing results also with the Lookup-Plus-Linear model. The data shows substantial reduction in both ΔXCEN and ΔYCEN, suggesting that there has been substantial drift over time. Nonetheless, it still shows the variation in ΔYCEN that correlates with the *Hinode* eclipse season.

For validation, we aligned *Hinode*/SOT-SP data from 2006 – 2010 with *SOHO*/MDI (Scherrer et al. 1995), using the MDI `mdi.fd_M_96m_lev182` data series. We took all of the corresponding *Hinode*/SOT-SP data from 2006 – 2010 meeting our criteria and used the procedure described in Section 3 to co-align it with *SOHO*/MDI data. Since *SOHO*/MDI is a LOS magnetogram, it was necessary to change the quantities used in alignment. We matched the magnitude of the *SOHO*/MDI LOS field and with *Hinode*/SOT-SP field strength and polarization. Additionally, due to the large change in scale between the two instruments, we found fewer inliers, and so to maximize the number of available correspondences, we used the full *Hinode*/SOT-SP scan rather than the filtering by time. We plot the results in Figure 15. Among scans within a year of launch, the average ΔXCEN is $1.8''$, and the average ΔYCEN is $17.3''$. The larger ΔYCEN is
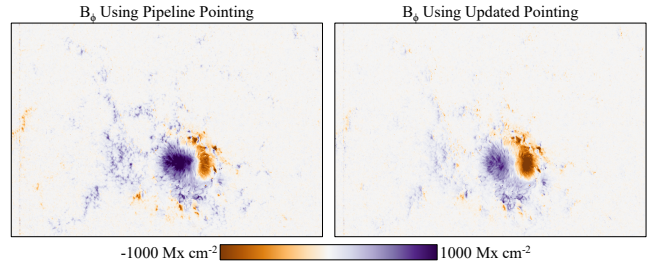

Figure 16. Impact of substantially incorrect `XCEN` **in the headers.** A small fraction of the *Hinode*/SOT-SP scans have pointing **in the header** that is substantially incorrect (by hundreds of arcseconds). Our previous figures have reported results with an upgraded `XCEN` calculated from the coordinates. One scan with an incorrect header appears on 2019 May 07, 06:49:34 TAI (the same AR, but not same scan as shown in Figure 17). We plot B_ϕ calculated with both the pipeline header pointing (`XCEN` = -339.6″) and our updated pointing (`XCEN` = -586.7″). The substantial change dramatically changes the configuration. For reference, a more typical pointing update occurs a few hours later, at 08:39:32. In this more typical update, the pipeline pointing is `XCEN` = 614.7″ and our updated pointing is `XCEN` = 589.1″.

driven by the eclipse-season jump: the averge ΔYCEN in the first six months is just $1.65''$. The alignments that are fit to *SOHO*/MDI continue to show the trends of Figure 10, and the prediction of the model fit later in the mission produces a reasonable extrapolation to the mission launch time.

6. PHYSICALLY-RELEVANT IMPLICATIONS

The corrections to the keywords `XCEN`, `YCEN` are typically on the order of 30 arcsec. The impact of this is different depending on where on the disk *Hinode*/SOT-SP was observing. Near disk center, a shift of 30 arcsec in `XCEN`, for example, corresponds to a change in longitude of less than $2°$, while the same change in XCEN for an observation nominally at a longitude of $70°$ is $\sim6°$, and increases further approaching the limb.

We note that some reported `XCEN`, `YCEN` are wrong due to suboptimal in-pipeline calculations. These can dramatically change the interpretation of regions, as shown for B_ϕ in Figure 16. In our analysis of the updates to the pointing in Section 5, we ignored these by re-calculating `XCEN`, and in this section, we exclude such data from our analysis to avoid conflating the dozen-of-arcsecond changes impacting most of the data with the hundreds-of-arcsecond changes impacting some of the data that can be fixed by a quick re-calculation.

Changes even at the scale of dozens of arcseconds, have physical implications, especially closer to the limb. We analyzed two active regions acquired on 2019 May
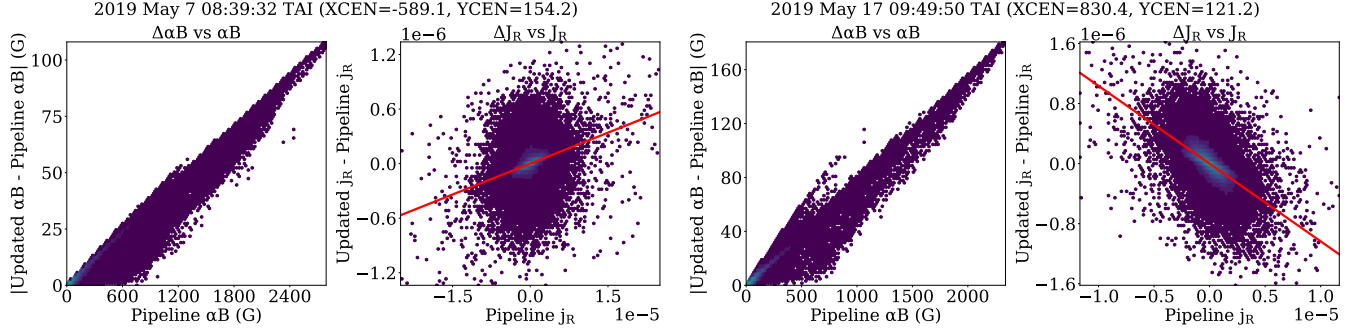
Figure 17. Analysis of how updating the pointing information changes physical quantities. We analyze two *Hinode*/SOT-SP scans from early May 2019 on opposite sides of the Sun's disk (left: 2019 May 07 08:39:32 TAI at `XCEN`=-589.1; right: 2019 May 17 09:49:50 TAI at `XCEN`=830.4) with `XCEN` determined via alignment. In each case, we plot the magnitude of the vector change in heliographic components against the magnitude on left and the change in radial current density against radial current density on right (along with a line of best-fit in **red**). Updating the pointing information changes the vector moderately (≈5%) and the radial current density more (≈10%). Moreover, the update to the radial current density is not consistent as seen by the skew in Δj_R-vs-j_R. On the the left panel's case, negative values of j_R are changed to be more negative and positive values are changed to be more positive. On the right panel's case, this trend is reversed: negative values are made more positive positive and positive values are made less positive. This trend holds true over larger amounts of data. Density legend: white: not observed; min ▬▬▬▬▬ max.

7 and 2019 May 17. Given the updated coordinates, one can re-compute the heliographic components of the field as well as other quantities such as the radial current density (j_R), which is important for determining the free energy in the corona. We plot the magnitude of the vector difference in $\alpha\mathbf{B}$ (i.e., $||[\alpha B_\phi - \alpha B'_\phi, \alpha B_\theta - \alpha B'_\theta, \alpha B_R - \alpha B'_R]||$) against the magnitude $||\mathbf{B}||$, as well as the signed change in j_R (i.e., $j'_R - j_R$) against j_R. We show density plots in Figure 17 for each active region. The vector difference has a magnitude of ≈5% while the radial current density changes by ≈10%. This difference occurs because, in addition to the heliographic components of the field, the distance between the pixels is changed. The change in inter-pixel distance alters the finite differences used to determine the radial current density. The change in radial current density, moreover, is not uniformly spread across the scan but instead appears as spatially distinct patterns as shown in Figure 18(left).

The positive bias of the corrections leads to bias in estimates of radial current density. On one limb, the updated inter-pixel distances always become smaller and on the other limb, they always become larger, resulting in the skew shown in Figure 17(right). We investigate this further with a set of 126 random scans from the dataset that: (a) have $-200'' <$ `YCEN` $< 200''$, to avoid `YCEN` to be a confounding factor; (b) have Δ`XCEN` > 10 so as to avoid small changes (although most scans satisfy this); (c) do not have *enormous* changes in pointing (Δ`XCEN` $> 100''$ or Δ`YCEN` $> 100''$) that are due to a miscoding of `XCEN` from the pipeline pointing. For each, we re-compute the radial current density j'_R and compare it

with the radial current density j_R as computed with the pipeline pointing. We then compute the optimal linear scaling from the pipeline to updated quantities (including an offset term), and plot this scaling as a function of disk location. As can be seen in Figure 18(right), there is a strong correlation between disk location and the scaling. This means that regions observed close to the west limb will have a total current which is underestimated when using the original pointing information, while those observed close to the east limb will have the total current overestimated.

We present here an illustration only of how recomputing the heliographic components of the field with the updated pointing changes the results, but the heliographic components of the field rely on a disambiguation algorithm that uses the radial component of the current density to determine the most likely solution. Thus the disambiguation can also be impacted by the change in the pointing, leading to even large changes in the heliographic components than shown here.

7. DISCUSSION AND CONCLUSIONS

This paper has presented an approach for aligning and thus spatially calibrating Level 2 *Hinode*/SOT-SP data and *SDO*/HMI data. This is challenging, not only because *Hinode*/SOT-SP is a scanning slit spectrograph but also because the instruments have varying characteristics. Indeed, as Sainz Dalda (2017) and Higgins et al. (2022) point out, the magnetograms produced by the pipelines of the two instruments used in this study differ not only in spatial resolution but also key implementation details. These differences result in systematic
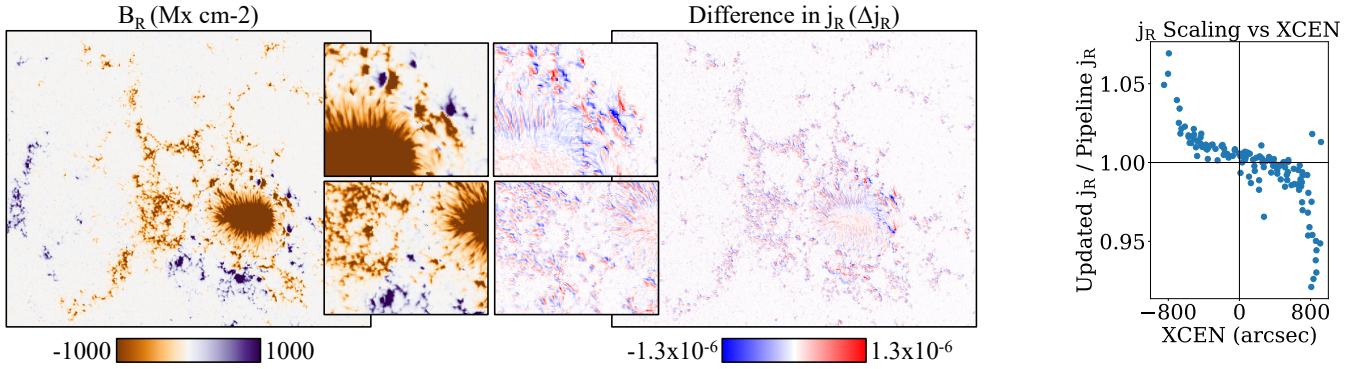
Figure 18. Illustration of the non-uniform changes to j_R. *Left two panels*: we plot B_R for context and the difference in j_R, along with zoom-ins in between the images. The changes in j_R are *not* spatially uniform, but exhibit distinct spatial trends consisting of ridges. *Right panel:* The overall relative scaling between j_R calculated with the updated pointing information compared to j_R calculated with the pipeline pointing information. On the left hemisphere, the new j_R is almost always bigger; on the right, it is almost always smaller.

differences in structures such as plage (Griñón-Marín et al. 2021; Leka et al. 2022).

Nonetheless, the approach presented does well at alignment and enables backing out both scan-invariant parameters in the form of relative scaling as well as scan-variant parameters in the form of pointing and rotation. The method's success is built on well-understood techniques from computer vision of correspondence and geometric model fitting that generalize well and have formed the backbone of many of computer vision's pre-deep learning successes. We do stress, though, that proper accounting for *Hinode*/SOT-SP's scanning slit mechanism is critical for the success of this method, and that while the *Hinode*/SOT-SP Level 2 data is presented in an image-like format, it cannot be directly treated as such for alignment. Earlier versions of this study empirically found noticeably worse performance when slit position was not accounted for. This need to properly understand instruments underscores the need for careful, sustained interdisciplinary collaboration in order to obtain correct results. Accounting for all the particulars of the instrument separates this works' alignment from our earlier work (Higgins et al. 2022) and reduces accurate alignment's dependence on post-processing steps such as optical flow. However, we do note that pixel-perfect alignment of all parts of the image (e.g., for the creation of synthetic magnetogram training) would require an additional optical flow step due to the non-uniform evolution and flow of the Sun.

We note that the techniques applied in this paper may help in other challenging solar co-alignment settings. One appealing aspect of the technique presented is that correspondence extraction is robust to substantial differences between the data that are being aligned, as illustrated by the alignment of field

strength from *Hinode*/SOT-SP and flux density from *SDO*/HMI (which are different quantities) and inclination from *Hinode*/SOT-SP and *SDO*/HMI (which are different in each instrument's data due to the treatment of fill factor, as noted by Sainz Dalda (2017)) and Leka et al. (2022). Moreover, as explored in this work, the use of point correspondences enables the seamless use of multiple modalities via the concatenation of correspondences as well as the handling of temporal evolution by the removal of correspondences between pixels taken at substantially different times. One benefit for large-scale analysis not demonstrated in this work is that the outlier-robustness makes this technique particularly effective even in the presence of large corrupted regions. In the setting for which these techniques achieved their first success (large-scale reconstruction Snavely et al. (2007); Schönberger & Frahm (2016)), this robustness enabled them to reconstruct large-scale 3D models of landmarks in the presence of tourists, vehicles, and changing seasons.

If the *SDO*/HMI scaling is correct, our experiments across thousands of scans spanning nearly a decade of the *Hinode* mission suggest that the *Hinode*/SOT-SP scaling is smaller than the values reported in the Level 2 headers and close to the nominal Level 1 values. This result contradicts the results found by Centeno et al. (2009), who cross-calibrated *Hinode*/SOT-SP with *Hinode*/SOT-FG BFI, which also observes a limited field of view (218″ by 109″) using the same base telescope. In the intervening time, the community has gained access to *SDO*/HMI, which observes the full ≈2000″-wide solar disk, albeit from a different vantage point.

Our procedure can only provide information about *relative* scaling, and cannot conclusively answer questions about *absolute* scaling. However, *SDO*/HMI

has several advantages over *Hinode*/SOT-FG BFI and *Hinode*/SOT-SP for serving as the source of reference measurements. First, *SDO*/HMI's full disk observations show the limb of the Sun. This provides strong constraints on scaling compared to the narrower fields of view on active regions observed by *Hinode*/SOT. As a secondary consideration, *SDO*/HMI's observations function far more like a global shutter (i.e., capturing a consistent snapshot of the Sun at a particular time): the Stokes vector observations used for HMI magnetograms are measured in approximately 2 minutes (although the data series we use averages multiple aligned observations over 12 minutes to reduce noise). In contrast, *Hinode*/SOT-SP's use of a scanning slit means that rightmost pixels are acquired dozens of minutes later than the left ones. These complexities in the spatial dimension are necessary for *Hinode*/SOT-SP's far higher spectral resolution and sampling, but they make *SDO*/HMI a more compelling global reference instrument.

The translations we fit with our model provided updated estimates of the pointing information of *Hinode*/SOT-SP for over twelve thousand scans spanning a decade. Currently, this pointing information is obtained by pointing obtained from other instruments aboard *Hinode*. As we showed, the errors in the current pointing information leads to incorrect estimates of physical quantities, with the particular example of radial current density. Our results suggest that the relative pointing calibration has both drifted slowly over time and that changes in temperature cause substantial seasonal changes in pointing around *Hinode*/SOT-SP's eclipse season (and not during *SDO*'s eclipse seasons, which occur at different times of the year). These changes in pointing are well predicted by temperature onboard *Hinode*, suggesting that thermal expansion is the underlying cause.

Similar pointing corrections were reported for *Hinode*/EIS by (Mariska 2016; The Hinode Review Team et al. 2019), who obtained these corrections by cross-calibrating *SDO*/AIA using different methods. Our results show similar issues on *Hinode*/SOT-SP with similar scales and help narrow down the cause. Given that *Hinode*/SOT-SP and *Hinode*/EIS do not share optical pathways and have different pixel sizes, it seems less likely that internal calibration for the instruments is shifting identically with precisely the right size. Instead, it seems more likely there is a single shift in *Hinode*'s attitude and orbital control system.

These results underscore the value of having multiple instruments with varying capabilities: the pointing issue would be difficult to identify and explain without *SDO*. The co-observation of *SDO* further enable us to extract data: our updated estimates of scaling and pointing information for *Hinode*/SOT-SP opens the door to a larger and more precisely aligned version of the joint dataset that powered the SynthIA system developed by Higgins et al. (2022).

REFERENCES

Borrero, J., Tomczyk, S., Kubo, M., et al. 2011, Solar Physics, 273, 267

Brown, M., & Lowe, D. G. 2003, in ICCV

Centeno, R., Lites, B., de Wijn, A., & Elmore, D. 2009, arXiv preprint arXiv:0902.0027

Centeno, R., Schou, J., Hayashi, K., et al. 2014, Solar Physics, 289, 3531, doi: 10.1007/s11207-014-0497-7

Civera, J., Grasa, O. G., Davison, A. J., & Montiel, J. M. 2010, Journal of field robotics, 27, 609

Community Spectropolarimetric Analysis Center (CSAC). 2006, doi: 10.5065/D6JH3J8D

Fischler, M. A., & Bolles, R. C. 1981, Communications of the ACM

Grec, C., Uitenbroek, H., Faurobert, M., & Aime, C. 2010, Astronomy & Astrophysics, 514, A91

Griñón-Marín, A. B., Pastor Yabar, A., Liu, Y., Hoeksema, J. T., & Norton, A. 2021, The Atrophysical Journal, 923, 84, doi: 10.3847/1538-4357/ac2aa8

Hartley, R. I., & Zisserman, A. 2004, Multiple View Geometry in Computer Vision, 2nd edn. (Cambridge University Press, ISBN: 0521540518)

Higgins, R. E. L., Fouhey, D. F., Antiochos, S. K., et al. 2022, The Astrophysical Journal Supplement, In Press

Hoeksema, J. T., Liu, Y., Hayashi, K., et al. 2014, Solar Physics, 289, 3483

Hurlburt, N., Slater, G., Tarbell, T., Berger, T., & Katsukawa, Y. 2009, Hinode Solar Optical Telescope Data Analysis Guide, SOT Data Guide

Kosugi, T., Matsuzaki, K., Sakao, T., et al. 2007, The Hinode Mission, 5

Leka, K. D., Wagner, E. L., Griñón-Marín, A. B., Bommier, V., & Higgins, R. E. L. 2022, SoPh, 297, 121, doi: 10.1007/s11207-022-02039-9

Levenberg, K. 1944, Quarterly of applied mathematics, 2, 164

Lites, B., Casini, R., Garcia, J., & Socas-Navarro, H. 2006, Memorie della Società Astronomica Italiana, 78

Lites, B., Akin, D., Card, G., et al. 2013, Solar Physics, 283, 579

Lowe, D. 2004, IJCV, 60, 91

Mariska, J. 2016, EIS/AIA Coalignment, EIS Software Notes

Marquardt, D. W. 1963, Journal of the society for Industrial and Applied Mathematics, 11, 431

Mikolajczyk, K., & Schmid, C. 2005, IEEE transactions on pattern analysis and machine intelligence, 27, 1615

Mur-Artal, R., & Tardós, J. D. 2017, IEEE transactions on robotics, 33, 1255

Norton, A., Graham, J. P., Ulrich, R., et al. 2006, Solar Physics, 239, 69

Pesnell, W. D., Thompson, B. J., & Chamberlin, P. C. 2012, SoPh, 275, 3, doi: 10.1007/s11207-011-9841-3

Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. 2011, in International conference on computer vision, 2564–2571

Sainz Dalda, A. 2017, The Astrophysical Journal, 851, 111

Scherrer, P., Bogart, R., Bush, R., et al. 1995, Solar Physics, 162, 129

Schönberger, J. L., & Frahm, J.-M. 2016, in Conference on Computer Vision and Pattern Recognition (CVPR)

Schou, J., Scherrer, P. H., Bush, R. I., et al. 2012, SoPh, 275, 229, doi: 10.1007/s11207-011-9842-2

Snavely, N., Seitz, S., & Szeliski, R. 2007, International Journal of Computer Vision

Szeliski, R. 2010, Computer vision: algorithms and applications (Springer Science & Business Media)

The Hinode Review Team, Al-Janabi, K., Antolin, P., et al. 2019, Publications of the Astronomical Society of Japan, 71, R1

The SunPy Community, Barnes, W. T., Bobra, M. G., et al. 2020, The Astrophysical Journal, 890, 68

Tsuneta, S., Ichimoto, K., Katsukawa, Y., et al. 2008, Solar Physics, 249, 167