Neural calibration of imaging Stokes polarimeters

F.A. Iglesias[1,2], A. Asensio Ramos[3,4], M. Sanchez[1] & A. Feller[5]

[1] *Grupo de Estudios en Heliofísica de Mendoza, Universidad de Mendoza, Argentina*
[2] *Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina*
[3] *Instituto de Astrofísica de Canarias, España*
[4] *Departamento de Astrofísica, Universidad de La Laguna, España*
[5] *MPI für Sonnensystemforschung, Alemania*

Received: 09 February 2024 / Accepted: 21 June 2024
©The Authors 2024

Resumen / Las técnicas actuales de calibración polarimétrica derivan la matriz de modulación del instrumento ajustando numéricamente un modelo instrumental a las mediciones de un conjunto de vectores de Stokes de calibración. Estas técnicas suelen estar limitadas a un error en la recuperación de los parámetros Stokes Q, U y V normalizados, en el rango de 1×10^{-2} a 1×10^{-3}. Este error suele aumentar cuando la respuesta del instrumento varía considerablemente a lo largo de su campo de visión y/o cuando existen efectos instrumentales que no están incluidos en el modelo de calibración asumido, como la no linealidad de la cámara. En este trabajo, proponemos una nueva técnica para calibrar polarímetros basada en un modelo compuesto por redes neuronales multicapa (NN). Este modelo está entrenado para aprender la matriz de modulación del instrumento dada una serie de parámetros de entrada, utilizando los mismos datos de calibración que se adquieren para las técnicas actuales. La principal ventaja de nuestro enfoque basado en NN es su flexibilidad para incorporar efectos instrumentales para los que no se dispone de un modelo preciso y, posiblemente mediante la fusión de datos de otros tipos de calibraciones relevantes, obtener una respuesta del instrumento más precisa. Presentamos un resultado preliminar del rendimiento de nuestro modelo utilizando datos sintéticos, donde se conoce la verdadera respuesta del instrumento.

Abstract / Current polarimetric calibration techniques derive the instrument modulation matrix by numerically fitting an instrumental model to measurements of a set of calibration Stokes vectors. These techniques are typically limited to an error on the retrieved normalized Stokes Q, U and V parameters, in the 1×10^{-2} to 1×10^{-3} range. This error commonly increases when the instrument response varies considerably across its field of view and/or when instrumental effects are present, which are not included in the assumed calibration model, such as camera non-linearity. We propose a new technique to calibrate Stokes polarimeters based on a model composed of fully-connected, multi-layer neural networks (NN's). This model is trained to learn the instrument modulation matrix given a set of input parameters, using the same calibration data that is acquired for the current techniques. The main advantage of our NN-based approach is its flexibility to incorporate instrumental effects for which no accurate model is available and, possibly through the fusion of data from other types of relevant calibrations, obtain a more accurate instrument response. We present a preliminary result of our model performance using synthetic data, where the ground truth is known.

Keywords / instrumentation: polarimeters — Sun: magnetic fields — instrumentation: high angular resolution

1. Introduction

During polarimetric calibration, we aim at obtaining the instrument transfer function, commonly given by its polarimetric modulation matrix (O), such that the modulated intensities ($\hat{I}_l$) registered for an input Stokes vector ($\hat{S}_l$) are: $\hat{I}_l = O S_l$. The matrix O is of dimensions $4xM$ for a polarimeter with M modulation states and, once measured ($\hat{O}$), it can be used to demodulate science data by doing $\hat{S}_l = \hat{O}^{-1}\hat{I}_l$. During calibration, a polarization state generator (PSG) is placed in front of the polarimeter to produce a set of different input Stokes vectors. The vector generated during the $l-th$ measurement can be written in matrix form as $S_{psg,l} = M_{psg}(\delta_l, \Phi_l, \Theta_l)[I_{src,l}\,0\,0\,0]^T$, where $M_{psg}(\delta_l, \Phi_l, \Theta_l)$ and $I_{src,l}$ are the PSG Mueller matrix, PSG retarder retardance and position angle, the PSG polarizer position angle, and the PSG input intensity level, respectively. Multiple calibration inputs are generated with the PSG to form a set of simultaneous equations. The current approaches derive an estimation of O, $\hat{O}$, from such a calibration data via Least Squares Minimization (LSM). If only $\hat{O}$ is a free parameter, the problem is linear and can be solved using the pseudo-inverse of $\hat{I}S_{psg}$ (Collados, 1999). If one also include as unknowns δ_l, Φ_l, and $I_{src,l}$, the problem becomes non-linear and is solved using an appropriate minimization algorithm (e.g. Levenberg-Marquardt). The current approaches are:

a) Modeled modulation matrix (Collados, 1999): The matrix $\hat{O}$ is modeled with few free parameters that are found, possibly along with parameters in M_{psg}, by LSM of $E_{s1} = \hat{O}^{-1}\hat{I}_l - S_{psg,l}$ or $E_I = \hat{O}^{-1} - S_{psg,l}$.

✉ *Contact* / franciscoaiglesias@gmail.com Research article

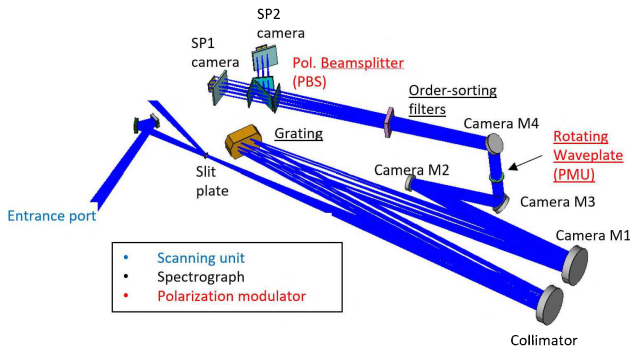

Fig. 1. Optical layout of SUSI narrow-band channel

b) Modulation matrix (Beck et al., 2005): All the $4xM$ elements of $\hat{O}$ (or $\hat{O}^{-1}$, possibly along with parameters in M_{psg}, are found by LSM of E_{s1} or $E_I h$. This is the most widespread method.

c) Response Matrix (Ichimoto et al., 2008): A known theoretical demodulation matrix O^{-1}_{mod} is used to fit only the $4x4$ elements of a response matrix, $\hat{X}$, by LSM of $E_{S2} = \hat{X} O^{-1}_{mod} \hat{I}_l - S_{psg,l}$.

The fits are typically done independently per pixel using binned calibration data, to reduce the influence of noise and still fit the variation of $\hat{O}$ across the instrument Field of View (FOV). The main limitations of these approach are:

i. In case of δ_l, Φ_l, and $I_{src,l}$ vary across the FOV, the PSG rotation couples the calibration equations of the different pixels. Performing the fit of all the pixels simultaneously is computationally very expensive for large size detectors.

ii. The calibration model and the scaling of the computational complexity do not favour the simultaneous fitting of various data sources that may benefit from it. For example, in a dual beam polarimeter, the correct approach is to simultaneously fit a single M_{psg} and two $\hat{O}$, one per channel.

iii. The minimum values reached on the normalized E_{S1} (or E_{S2}) are limited to the 10^{-2} to 10^{-3} range, even when photon noise in the calibration measurements is an order of magnitude lower. This is most likely due to features and/or artefacts no considered in the polarization model, such as camera non-linearity. Including such effects via detail modeling is a laborious and heavily instrument dependent.

To tackle this limitations, we propose a new technique to calibrate imaging Stokes polarimeters based on a model composed of multi-layer neural networks (NN's).

2. Instrument and calibration measurements simulation

We developed a numeric simulation code (SUSIM) that models the polarimetric aspects of the Sunrise Ultraviolet Spectropolarimeter and Imager (SUSI, see Feller et al. 2020) using the Mueller matrices of all the optical components in the beam path. SUSIM also incorporates a simple camera model, including sensor bias, conversion gain, noise and non-linearity. The SUSIM model parameters can be set to vary across the FOV to study the polarimetric effects of various instrumental artifacts. SUSI is a scanning slit-spectrograph that includes a dual-beam polarimeter, see Fig. 1. We produced two types of synthetic data sets to assess the neural calibration accuracy and compare it to the calibration method (b):

- Polarimetric calibration: A total of 40 different $S_{psg,l}$ are generated for each calibration simulation, by changing the nominal PSG position angles, Φ_l, Θ_l. We use 10 values of $\Theta = \{0, 90, 135, 255\}$deg.

- Science measurements: We input to SUSIM synthetic full-Stokes images of the Solar photosphere obtained with the MSP/University of Chicago Radiactive MHD code (MURaM, see (Przybylski et al., 2022).

3. Neural calibration model

We model the main quantities involved in the calibration by implicit representations using simple NN's, wich are trained to learn the polarimeter $\hat{O}$ (NN_O), and errors in the PSG $\delta_l(\mathrm{NN}_\delta)$, $\Phi(\mathrm{NN}_\Phi)$, and $I_{src,l}(\mathrm{NN}_l)$, see Fig. 2. The input and outputs of NN_O are the pixels coordinates $\{x, y\}$ and the $4xM$ elements of $\hat{O}$, respectively. We have tested a more general model, where instead of NN_O there is a NN using realistic calibration data. This is because the PSG's used in practice produce only totally polarized light, limiting the training dataset to only a small subset of possible inputs, i.e., the surface of the Poincare sphere. The three NN's outputting quantities that vary with the PSG configuration have as input the measurement index, l, or the corresponding rotated pixel coordinates $\{x', y'\}$. We implemented our model in PyTorch, other details are specified in Fig. 2.

4. Preliminary results

As a basic test of the neural approach, we produce two simulations. First a calibration measurement simulation where the exact behavior of the PSG is known, thus, NN_δ, NN_Φ and NN_I are not included in the neural model. The obtained calibration errors are shown in Fig. 3 and 4, showing a similar performance of both methods. Second, we used the calibration results to demodulate simulated solar measurements. The known input of the solar measurement simulation is a synthetic Stokes spectrogram obtained with the MPS/University of Chicago Radiative MHD (MURAM) code, including both quiet-Sun and a pore region around the 365 nm

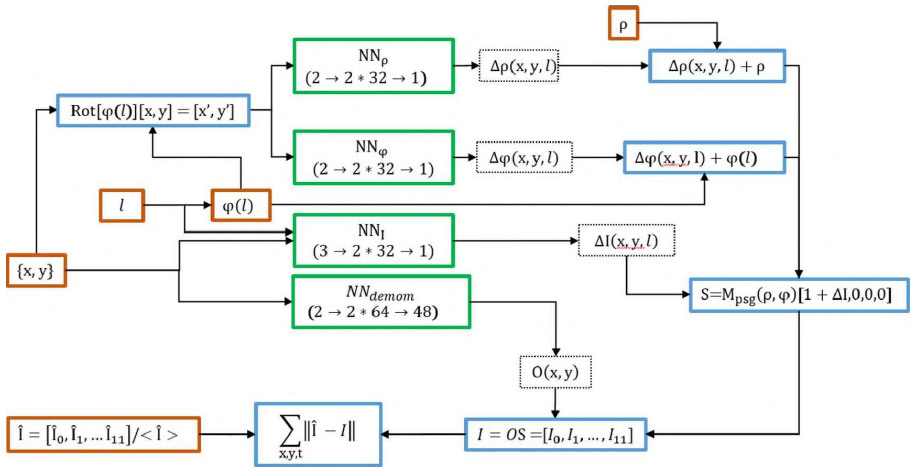

Fig. 2. Block diagram showing the data flow and main components of the proposed neural calibration model. Input data, NNs and computation steps are shown in orange, green, and blue boxes, respectively. The text inside the NN boxes specifies the input, hidden layer, and output dimensions. In this implementation, we use fully-connected, multi-layer perceptrons with ELU activation functions for all NNs. We train the model for ≈3k epochs, with each batch consisting of $8 \times 8\,\mathrm{px}^2$ patches taken randomly from input calibration images. Moreover, we adopt the Adam minimization algorithm and the CosineAnnealingRL scheduler for the learning rate.

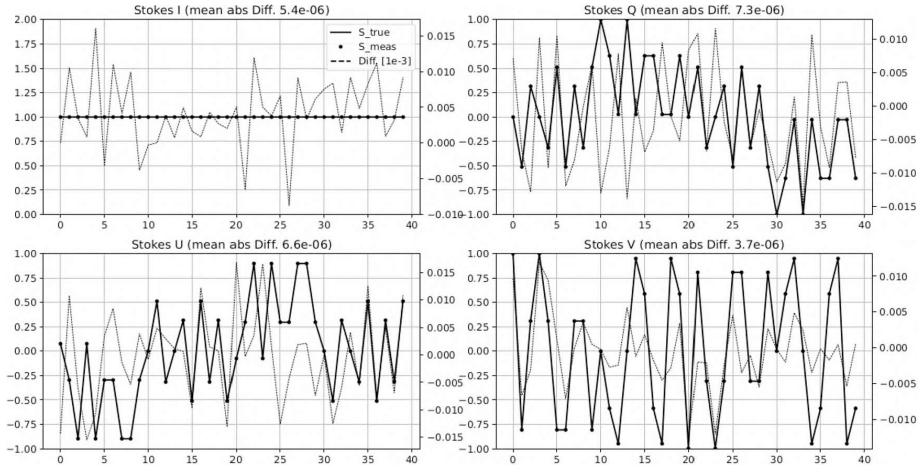

Fig. 3. True (S_{true}) and measured (S_{meas}) calibration Stokes parameters *(see the panels titles)* for method (b), along with their difference *(Diff., see the legend)*. For all quantities we plot the average over the image area for all the 40 calibration inputs (PSG position). Stokes Q, U and V are normalized with Stokes I and Stokes I is normalized to its value for the first calibration input.

region. The different between simulate and true solar signals are shown in Fig.5 and 6.

The future steps to improve the model flexibility to account for more realistic calibration scenarios are (a) Simulate unknowns in the PSG to assess NN_δ, NN_Φ and NN_I performance, (b) Study using penalties in the loss function to enforce the physicality of O, (d) Simulate a more realistic calibration, e.g., including camera NL and straylight, to evaluate the flexibility of the neural architecture to incorporate other sources of calibration data.

Acknowledgements: FAI and MS are supported by The Max Planck Partner Group between the University of Mendoza and the Max Planck Institute for Solar System Research in Germany.

References

Beck C., et al., 2005, Astronomy and Astrophysics, 437, 1159

Collados M., 1999, B. Schmieder, A. Hofmann, J. Staude (Eds.), *Third Advances in Solar Physics Euroconference: Magnetic Fields and Oscillation*, *ASP Conference Series*, vol. 184, 3–22

Feller A., et al., 2020, C.J. Evans, J.J. Bryant, K. Motohara (Eds.), *Ground-based and Airborne Instrumentation for Astronomy VIII*, vol. 11447, 2260 – 2267, SPIE

Ichimoto K., et al., 2008, Solar Phys., 249, 233

Przybylski D., et al., 2022, A&A, 664, A91

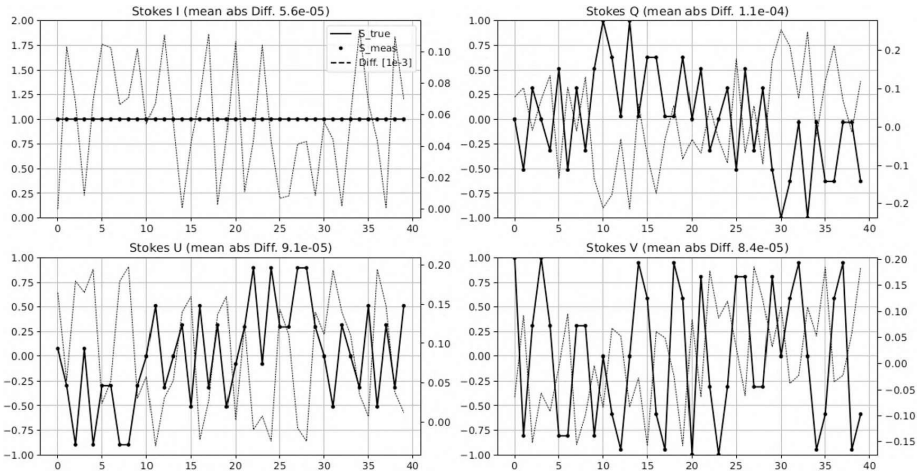

Fig. 4. Same as Fig. 3 for the neural calibration method. Note that the difference is larger than for method (b), however is well below the 1×10^{-3} achieved in real cases.

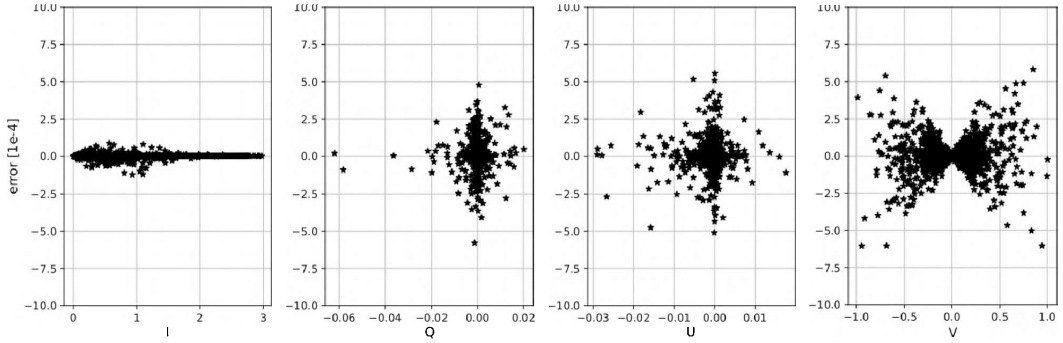

Fig. 5. Difference between true and measured Stokes parameters *(see horizontal axes label in each panel)* for the simulated solar measurement, versus Stokes signal level. These errors were obtained using calibration method (b).

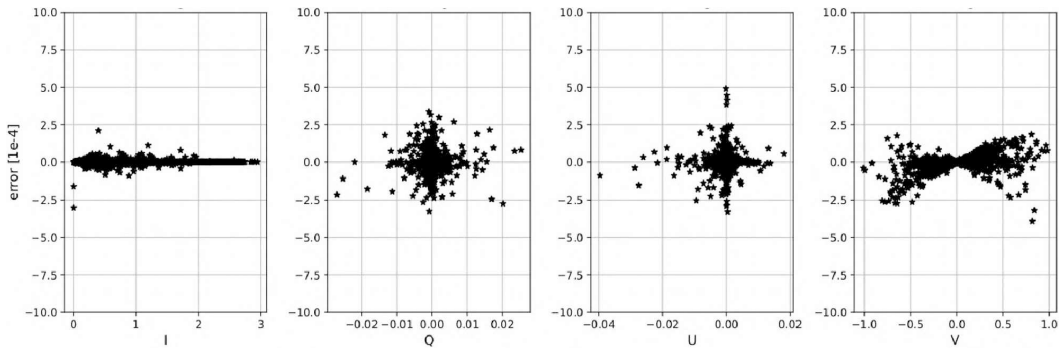

Fig. 6. Same as Fig. 5 for the neural calibration. The error levels are similar, to method (b), except for a slightly better performance in Stokes V. This, despite the neural approach having a larger calibration error, see Fig. 4.