ACTA IMEKO
ISSN: 2221-870X
September 2024, Volume 13, Number 3, 1 - 6

Automating flowmeter calibration process: Digital measurements from numerical displays using open-source optical character recognition tools

Gustavo Esteves Coelho[1], Alexandre Pinheiro[1], Álvaro Silva Ribeiro[1], Catarina Simões[1], Luís Lages Martins[1]

[1] *LNEC – National Laboratory for Civil Engineering, Av. do Brasil 101, Lisbon, Portugal*

ABSTRACT
This paper presents a methodology for obtaining digital machine-readable measurements from numerical displays images. The proposed method provides means to digitalize an automate a previously manual and labour-intensive laboratory procedure for flowmeters calibration. The proposed method allows to obtain machine-readable readings from remote numerical displays with available-off-the-shelf hardware and open-source software. By using smartphones for remote image capture and streaming and the Tesseract open-source OCR engine, is possible to leverage the infrastructure's digital transition, improve procedures efficiency and effectiveness while promoting sustainable actions with cost reductions.

Section: RESEARCH PAPER

Keywords: Metrology; automation; process efficiency; machine-readable data; open-source programming

Citation: G. Esteves Coelho, A. Pinheiro, Á. Silva Ribeiro, C. Simões, L. Lages Martins, Automating flowmeter calibration process: Digital measurements from numerical displays using open-source optical character recognition tools, Acta IMEKO, vol. 13 (2024) no. 3, pp. 1-6. DOI: 10.21014/actaimeko.v13i3.1767

Section Editor: Marija Cundeva-Blajer, Ss. Cyril and Methodius University in Skopje, North Macedonia

Received February 1, 2024; **In final form** June 4, 2024; **Published** September 2024

Corresponding author: Gustavo Esteves Coelho, e-mail: gfcoelho@lnec.pt

1. INTRODUCTION

The Industrial 4.0 era and particularly the Industrial Internet of Things (IIoT) lead to an unprecedented ability to use different types of sensors and generate enormous amount of measurement data [1], requiring traceability [2]. The heterogeneous nature of the equipment and the laboratory with inadequate Information Technology (IT) penetration, are often in the origin of the bottleneck for a fully digitalized calibration process.

The development and establishment of digital processes and digital infrastructures offers enormous potential for overall calibration process efficiency [3]. However, this process is often doomed to manual interaction and is usually labour-intensive. Nevertheless, it is possible to reduce human interaction and improve process efficiency and effectiveness by increasing the infrastructure's IT penetration, towards a fully digital and machine-readable calibration process.

The current accessibility to technology combined with the maturity of open-source software, including Artificial Intelligence (AI), allows the introduction and interoperability of digital processes with available off-the-shelf hardware and custom software, with reduced financial investments. The method proposed for the calibration of flowmeters with numerical displays, aims to aggregate and automate in a single computer all the data collection and storage of a typical decentralized and manual calibration process, gathering images remotely with cameras (image time synchronization is also discussed), followed by pre-processing the images and obtaining the machine-readable readings from the numerical displays by means of Optical Character Recognition (OCR).

This method is scalable to virtually any number of cameras and can be applied to both laboratory calibration and *in-situ* measurement processes. The software is developed with open-software solutions implemented in Python, with OpenCV for image pre-processing and Tesseract for OCR engine implementation.

2. INFRASTRUCTURE DESCRIPTION

The Unit of Hydraulic Metrology (UHM) is a R&DI infrastructure jointly coordinated by the Department of Hydraulics and Environment (DHA) and the Scientific

Figure 1. UHM-LNEC infrastructure overview.

Instrumentation Centre (CIC) of Laboratory for Civil Engineering (LNEC), with competence and capabilities to develop research in hydrology and hydraulics and to provide traceability to instrumentation and systems applied in a wide range of measurement quantities, namely, flow rate (mass and volumetric), flow speed, volume, level, and precipitation.

The laboratory infrastructure (Figure 1) has several hydraulic test benches allowing to establish different conditions to obtain flow rate by the primary gravimetric measurement using two weighing platforms (reaching 3 ton and 30 ton of mass) and the measurement of time using universal time counters, all traceable to primary standards of IPQ (Portuguese Institute for Quality, the Portuguese National Metrology Institute). The main experimental facility has the following operational capabilities:

- volumetric flow rate ≤ 0.500 m³/s;
- mass flow rate ≤ 400 kg/s;
- nominal diameter ≤ DN 400;
- maximum operating pressure ≤ 1.0 MPa;
- power ≤ 250 kW of electric power groups;
- power ≤ 75 kW for electric pumps not coupled to drive motors.

This Unit supports the skills that allow UHM-LNEC to be a Designated Institute for the flow rate and flow speed for liquids, according to the international recognition accepted by the BIPM in 2021 and confirmed by EURAMET in 2022. The management of UHM-LNEC is developed according to the LNEC Quality Management System complying with the requirements of the ISO/IEC 17025 standard [4].

The R&DI develops methods and apply processes to provide traceability and to perform metrological characterization related with several types of measuring instruments, namely, ultrasonic flowmeters, turbine meters, positive displacement flowmeters, differential pressure flowmeters, rotameters, mass flowmeters, Parshall flumes, among others [5]-[8].

This infrastructure has human resources with different academic backgrounds and technologies capable of promoting hydraulic metrology services and metrological information management, in a variety of areas of water resources management (water supply, undue inflow, agricultural uses and wastewater treatment), and in different frameworks (water management, industry, manufacturers, and customers).

UHM-LNEC operates in diverse areas, seeking to promote the quality of measurement in different areas of the economy, promoting measurement in the field, evaluating measurement and uncertainty in water supply networks for human consumption, in wastewater management systems, in water supply for industry [9] and agriculture, associating its action with consultancy, capacity building and decision-making processes. More recently, UHM-LNEC is developing skills for the measurement of precipitation, participating in projects that aim to monitor climate phenomena with a growing impact on society and find solutions to respond to the challenges of climate change [10].

3. THE DIGITAL TRANSITION CHALLENGE AND IMPLEMENTATION

The previously described infrastructure has been subject to successive modernization to foster and leverage the digital technologies and its transition. More recently, the infrastructure's automation legacy system was upgraded to comply with the framework of Industry 4.0 and towards the Industrial Internet of Things (IIoT) [11]. The integration of standard industrial communication protocols and modern technologies leverages the infrastructure's automation capabilities and big data analysis,

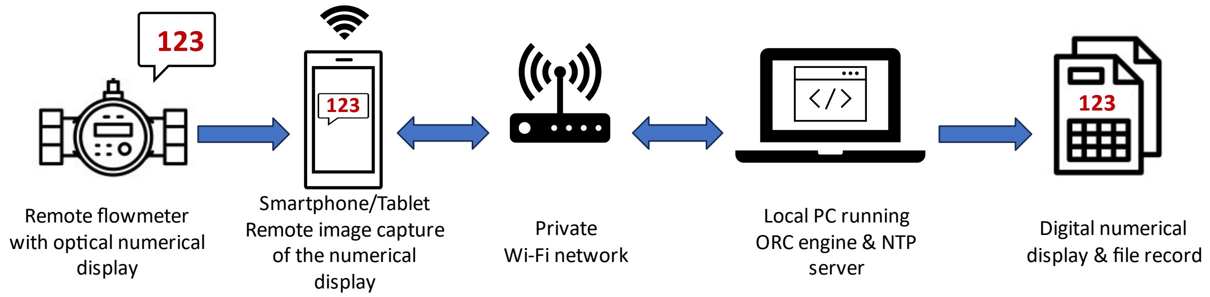


Figure 2. System implementation architecture.

aiming to improve efficiency and effectiveness of the calibration procedures, the auditing process, the infrastructure's maintenance and to comply with the quality management systems requirements for accreditation, including the management of resources, processes and risk assessment.

In the same perspective, the work presented here aims to contribute to the modernization of this infrastructure and to the ongoing digital transition, being the main goal to develop a solution for the acquisition of measurements obtained from flowmeters being calibrated in a machine-readable format, able to collect the raw measurement data directly in a digital media or support (e.g., standard XML file or spreadsheet). However, given the wide range of flowmeter devices that are currently calibrated at this infrastructure, showing different interfaces and numerical displays, a one-size-fits-all solution approach is difficult to reach. Additionally, the remote location of the equipment in the plant usually restricts the use of cables and wired devices to obtain the measurements of flowmeters under calibration. Nevertheless, the majority of the flowmeter's are at least equipped with one numerical display that enables the measurement by inspection.

To tackle these issues, the proposed approach general principle employs imaging devices to capture the remote images from the numerical displays in real-time and obtain the measurements in a numerical machine-readable format by means of Optical Character Recognition (OCR).

In Figure 2 the implemented system architecture is presented. From left to right:

- near the flowmeter under calibration an imaging device (such as smartphone) is mounted;
- the imaging device is positioned to obtain an image of the flowmeter's numerical display;
- the image is stream over Wi-Fi to a local PC; The PC implements the OCR engine and image processing and also provides Network Time Protocol (NTP) for clock synchronization between the capturing devices and the PC; and
- the OCR engine converts the numerical display from the image to machine-readable numbers format that can be stored in digital support such as XML or spreadsheets files, for further machine-to-machine processing, e.g., Digital Calibration Certificates (DCC) [12],[13].

4. EXPERIMENTAL SYSTEM IMPLEMENTATION

The remote flowmeter's numerical display can be capture with available off-the-shelf smartphones, considering that these devices have good image quality with low optical distortion and are usually Wi-Fi capable. The smartphone is mounted in a tripod with the camera properly facing the flowmeter's numerical display, as depicted in Figure 3 and Figure 4. The flowmeter's image is then streamed over Wi-Fi to a PC in the monitoring and command room (see Figure 5). There are several implementations and software available to stream the video from a smartphone to a PC. In this case we used the android IP Webcam app (available in Google Play), that provides unicast video streaming with session validation capabilities. Although, private Wi-Fi network is available at UHM-LNEC´s infrastructure, the later option is used for calibrations in field environments where private networks are not available, and communications security are mandatory.

Although the image processing is conducted in an asynchronous model by design, the NTP synchronization of the

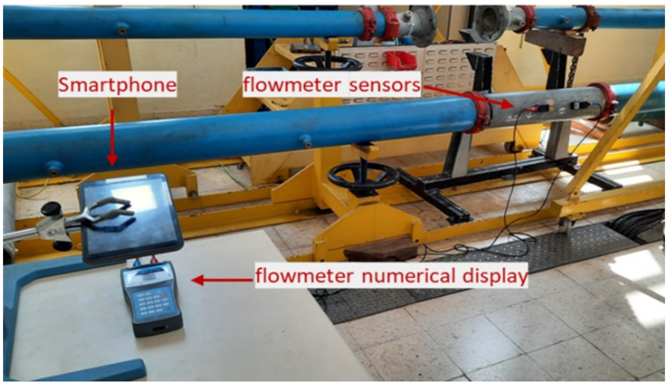


Figure 3. Experimental setup example for ultrasonic flowmeter calibration at the UHM-LNEC plant.

Figure 4. Close up of the flowmeter's numerical display and the smartphone for image streaming.

Figure 5. Monitoring and command room of the UHM-LNEC plant.

devices provides means of time synchronization between the flowmeter measurement (primary dada measurements) and the standard measurement and allows traceability and auditability of the overall process. For this purpose, the image streaming is time stamped locally in the smartphone device and the PC provides another time stamp for the measurements during the calibration process. In order to establish a time reference for the entire process, the capturing devices (smartphones) and the image processing unit (PC) are time synchronized with the same NTP server. This NTP server is provided locally from the PC to have the same time reference. The local NTP server is in turn synchronized downstream with a global NTP server, providing global time synchronization and time traceability to the overall system.

With the IP Webcam app installed in the smartphone, the video stream can be accessed through the smartphone's IP directly in a web browser. This configuration provides a real-time video stream and video feed of the remote flowmeter's numerical display to a PC located in our monitoring and command room, see Figure 5 and Figure 6.

The image processing application and OCR engine was developed in Python with fully open-source software: using OpenCV for the image pre-processing and using Tesseract for the OCR engine.

The first step is to is to create a bounding box around the area of numbers of interest in the image, namely in the area corresponding to the measurement of the flowmeter (see Figure 7, left). Since the OCR engine is sensitive to the image quality, e.g., light exposure, display contrast and especially the presence of light reflections in the display surface, the next step the numerical display image is converted in binary (black and white) video format (see Figure 7, right). This step is needed to achieve adequate confidence levels from the OCR results. With the binary image output the operator can evaluate the appropriate image quality to send to OCR engine.

With a single command in the PC, the binary image is sent to Tesseract OCR engine, where the machine-readable of the numerical display from the OCR engine is overlayed in the image along with the confidence value, as depicted in Figure 8, Figure 10 and Figure 12 in next section. All steps of the automation procedure are recorded in digital support in XML and spreadsheet files, respectively. From the numerical display's pictures to the binary images, including the digital information of the OCR's numerical display and the corresponding confidence values, along with the PC time stamp, to the digital primary data measurements, for traceability and auditing purposes, thus complying with accreditation requirements for primary data.

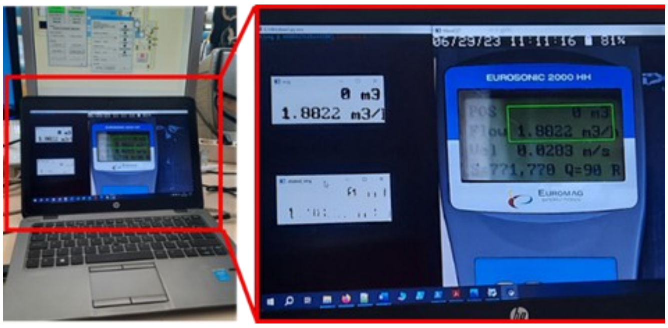


Figure 6. Remote image of the flowmeter display in the PC, in the monitoring and command room.

Figure 7. Left: Original image of the flowmeter numerical display with bounding box overlayed around the area of interest, Right: corresponding image converted in binary format.

5. RESULTS

To assess the proposed methodology, several video-feeds of the flowmeter's numerical displays were acquired in different angles and light conditions and evaluated with the OCR engine. In general, the results show that the implemented OCR engine is adequate to convert the numerical measurements from the numerical display to machine-readable format, to the majority of the flowmeters under calibration. The OCR results summary is presented in Table 1.

In Figure 8 an example of the numerical display of an ultrasonic flowmeter (model eurosonic 2000 HH) is presented. The OCR results are overlayed in the picture, as described in previous section, showing the obtained numerical conversion (left) and a confidence level of the OCR in % (right). Figure 9 shows the binary image output of the pre-processed image in the bounding box of Figure 8 before sending to the OCR engine to obtain the measurements in machine-readable format.

With proper display's contrast and light conditions, the OCR results in a correct numerical conversion. However, in the presence of low brightness and/or light reflections, the conversions results are inadequate as expected. This is an inevitable limitation of the OCR engine, nonetheless it can be mitigated by assuring a fair quality of the images before applying it to the OCR engine. To implement this, the image stream is

Figure 8. OCR result and confidence overlayed in the numerical display image.

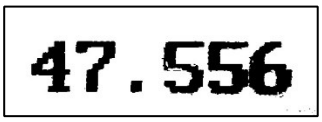
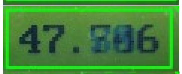
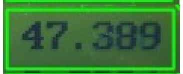
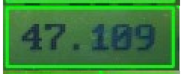


Figure 9. Binary image output of the pre-processed image of interest in Figure 8.

Table 1. OCR results summary.

Image Input	OCR output	Confidence score	Correct / Incorrect	Tested scenario
	47.556	93 %	Correct	Ideal conditions
	47.852	91 %	Correct	Ideal conditions
	47.746	87 %	Correct	Inadequate binary image
	47.389	83 %	Correct	Inadequate binary image
	47.169	73 %	Incorrect	OCR model limitation
	45	87 %	Incorrect	OCR model limitation
	14.138	89 %	Incorrect	Adverse light conditions

pre-processed to obtain the binary counter part of the images. Image binarization enables to obtain the best contrast between the background and the numbers to be converted and provide useful information to the operator about the quality of the image stream in the OCR's perspective.

Figure 10 shows an example of OCR output results in presence of adverse light conditions. In this case, although the image appears to be in fair conditions for human interpretation, the OCR engine outputs the incorrect numerical measurements when the binary image quality is inferior (see Figure 11).

Consequently, image pre-processing is also used in the early stage of the smartphone positioning in front of the numerical display of the flowmeter under calibration, to achieve the best possible quality of the image stream to the OCR engine.

Nevertheless, the OCR can also output incorrect values due to limitation in the OCR model. The case presented in Figure 12 is an example of the incorrect identification of the zero value. Although the correspondent binary image is acceptable according to Figure 13, the OCR results exhibit an incorrect value of the number zero in the flowmeter's display. This is a limitation of the current OCR engine for the slashed zero "Ø" that is incorrectly converted to number "6". This problem only

Figure 10. Influence of light reflections in the OCR output.

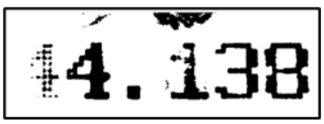


Figure 11. Binary image output in the presence of adverse light effects.

Figure 12. Slashed zero OCR engine problem.

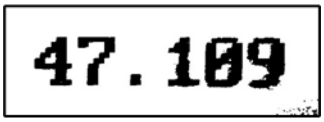


Figure 13. Corresponding binary image.

affects displays using slashed zero and is further discussed in section 6. However, the confidence low values give some insight of this limitation, and for this reason the OCR output is always evaluated and stored along with the confidence values to detect and potentially manage these situations.

6. DISCUSSION AND FUTURE WORK

A solution to digitalize and automate a manual and labour-intensive laboratory procedure was presented. With available off-the-shelf hardware (smartphone and laptop) and open-source software was possible to develop and implement a new tool that enables to read remote numerical displays of a flowmeter and obtain its measurements in a machine-readable format ready to store directly in a computer, with reduced human interaction.

The results showed that it is possible to improve laboratory procedure efficiency and effectiveness with minimal financial investment. The proposed solution also promotes sustainable practices, important in the digital transformation, by reducing intermediate processes that are usually supported by pen-and-paper methods.

Since the image streaming is implemented over standard ethernet protocols, this method is also scalable to any number of smartphones and feasible for calibration of multiple flowmeters simultaneously, depending only on the devices processing power running the OCR engine.

Although the laboratory procedure for flowmeter's calibration usually requires human interaction, there are some procedures that is desirable to automate to further increment the process efficiency. Since the process for the calibration of a flowmeter device involves gathering simultaneous and synchronous measurements form both the device under calibration (primary measurements) and the standard measurement, future work addresses further reduction in human interaction towards a fully automated procedure.

The obtained results shows that OCR's outputs are acceptable with confidence levels above 90 %. This suggests that further automation in the calibration process is feasible. For example, to assure the OCR optimal results, one potential improvement is to implement a threshold approach with confidence levels above 90 % in the processing of the flowmeter measurements. Additionally, by simultaneously obtaining the standard measurement (standard measurements are normally obtained directly in machine-readable format), this further reduces the human intervention in the calibration process.

To reduce the influence of adverse light conditions resulting in the poor OCR performance, one feasible solution is to use artificial and controlled light environment, combined with low reflective backgrounds.

The problem regarding the slashed zero mismatch is a well-know and documented issue in the Tesseract open-source community. To address this issue adequate OCR models are need. These models can be obtained directly from the open-source community or by training new OCR models with adequate training data. Nevertheless, as already stated, this limitation only affects the measurements with devices that employs this specific type of numerical font.

By combining off-the-shelf hardware with standard protocols and standard digital support (XML and spreadsheet files), the proposed method is easily implemented in the most calibration laboratories seeking to automate procedures and to foster the digital transition, with minimal financial investments. Finally, by storing digitally all steps of process, the calibration procedure is entirely traceable and auditable.

REFERENCES

[1] S. Munirathinam, Chapter Six – Industry 4.0: Industrial Internet of Things (IIOT), Editor(s): Pethuru Raj, Preetha Evangeline, Advances in Computers, Elsevier, Volume 117, Issue 1, 2020, pp. 129-164, ISSN 0065-2458, ISBN 9780128187562.

[2] M. G. Cox, P. M. Harris, Measurement uncertainty and traceability, Measurement Science and Technology, Volume 17, Number 3, 2006
DOI: 10.1088/0957-0233/17/3/S13

[3] S. Eichstädt, Metrology for the digital age. Metrological Infrastructure, edited by Beat Jeckelmann and Robert Edelmaier, Berlin, Boston: De Gruyter Oldenbourg, 2023, pp. 131-154.
DOI: 10.1515/9783110715835-006

[4] A. S. Ribeiro, L. Dias, A. C. Martins, H. Dimas, R. Mendes, Flow measurement traceability used to improve water management, Test & Measurement 2018 International Conference. NLA – South Africa, Western Cape (South Africa), 2018.

[5] M. A. Silva, C. Amado, A. S. Ribeiro, D. Loureiro, Uncertainty evaluation in time-dependent measurements, Measurement 196, 2022.
DOI: 10.1016/j.measurement.2022.111196

[6] A. Ribeiro, D. Loureiro, M. C. Almeida, M. Cox, J. A. Sousa, M. Silva, L. Martins, R. Brito, A. C. Soares, Uncertainty evaluation of totalization of flow and volume measurements in drinking water supply networks, FLOMEKO 2019 – 18th International Flow Measurement Conference, Lisboa (Portugal) , 26-28 June 2019. Online [Accessed 12 August 2024]
https://www.imeko.info/publications/tc9-2019/IMEKO-TC9-2019-093.pdf

[7] C. Simões, A. S. Ribeiro, M. C. Almeida, D. Covas, Uncertainty evaluation related with the fitting of probability distributions to rainfall experimental data, IMEKO XXIII World Congress - Measurement: sparking tomorrow's smart revolution, Yokohama, Japan, 30 August–3 September 2021.
DOI: 10.1016/j.measen.2021.100284

[8] R. S. Brito, M. C. Almeida, A. Ribeiro, Enhancing hydraulic data reliability in sewers, Water Practice & Technology Vol 17 No 1, 2022, 431.
DOI: 10.2166/wpt.2021.093

[9] C. Simões, A. S. Ribeiro, M. C. Almeida, D. Covas, Uncertainty analysis of average flow duration curves used in hydropower plants feasibility studies, IMEKO TC11 & TC24 Joint Hybrid Conference, Dubrovnik, Croatia, 17-19 October 2022.
DOI: 10.21014/tc11-2022.18

[10] C. Simões, A. S. Ribeiro, M. C. Almeida, D. Covas, L. Dias, G. E. Coelho, Validation of a laboratory method for the traceability of a rainfall weighing gauge, Acta IMEKO 13 (2024) 1, pp. 1-6.
DOI: 10.21014/actaimeko.v13i1.1322

[11] G. E. Coelho, A. S. Ribeiro, J. dos Reis, C. Simões, Industry 4.0 Legacy Systems Integration Case Study, IMEKO TC11 & TC24 Joint Hybrid Conference, Dubrovnik, Croatia, October 17-19, 2022.
DOI: 10.21014/tc11-2022.08

[12] S. G. Hackel, F. Härtig, J. Hornig, T., Wiedenhöfer, The digital calibration certifcate. PTB-Mitteilungen 127, 75–81, 2017. URL:https://oar.ptb.de/resources/show/10.7795/310.20170403
DOI: 10.7795/310.20170499 .

[13] S. Hackel, F. Härtig, T. Schrader, A. Scheibner, J. Loewe, L. Doering, B. Gloger, J. Jagieniak, D. Hutzschenreuter, G. Söylev-Öktem, The fundamental architecture of the DCC, Measurement: Sensors, Volume 18, 2021, 100354.
DOI: 10.1016/j.measen.2021.100354