View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by CONICET Digital

AGU PUBLICATIONS

Water Resources Research

TECHNICAL REPORTS: DATA

10.1002/2017WR020897

Hydrometric Data Rescue in the Paraná River Basin

Andrés Antico[1,2], **Ricardo O. Aguiar**[3,4], **and Mario L. Amsler**[1,2]

[1]National Scientific and Technical Research Council (CONICET), Argentina, [2]Facultad de Ingeniería y Ciencias Hídricas, Universidad Nacional del Litoral, Santa Fe, Santa Fe, Argentina, [3]Delegación Paraná Inferior, Dirección Nacional de Vías Navegables, Rosario, Santa Fe, Argentina, [4]Facultad Regional Rosario, Universidad Tecnológica Nacional, Rosario, Santa Fe, Argentina

Key Points:

- We rescued historical hydrometric data of the Paraná River, one of the ten largest rivers in the world
- We offer the longest (last 143 years) continuous hydrometric records of the Paraná River
- The rescued data are valuable for studies of global and South American hydroclimate changes

Correspondence to:

A. Antico,
aantico@santafe-conicet.gov.ar

Citation:

Antico, A., Aguiar, R. O., & Amsler, M. L. (2018). Hydrometric data rescue in the Paraná River basin. *Water Resources Research*, *54*, 1368–1381. https://doi.org/10.1002/2017WR020897

Received 5 APR 2017
Accepted 2 NOV 2018
Accepted article online 8 JAN 2018
Published online 2 FEB 2018
Corrected 15 FEB 2018

This article was corrected on 16 FEB 2018. See the end of the full text for details.

Abstract The Paraná River streamflow is the third largest in South America and the sixth largest in the world. Thus, preserving historical Paraná hydrometric data is relevant for understanding South American and global hydroclimate changes. In this work, we rescued paper format data of daily Paraná water level observations taken uninterruptedly at Rosario City, Argentina, from January 1875 to present. The rescue consisted of the following activities: (i) imaging and digitization of paper format data, (ii) application of quality checks and homogeneity tests to the digitized water levels, and (iii) consideration of errors caused by gauge sinkings that may have occurred from 1875 to 1908. In addition, a rating curve was obtained for Rosario and it was used to convert water levels into discharges. The rescued water level observations and their associated discharge data provide the longest (last 143 years) continuous hydrometric records of the Paraná basin. The usefulness of these records was demonstrated by showing that the Paraná-Pacific Ocean links observed after 1900 in previous studies are also evidenced in our nineteenth-century discharge data. That is, high Paraná discharges coincided with El Niño events and with El Niño-like states of the Interdecadal Pacific Oscillation (IPO), whereas low discharges coincided with La Niña events and with La Niña-like IPO states.

1. Introduction

Historical hydrometric river data (e.g., water level observations) are of paramount importance for understanding past and present hydroclimate variability, predicting future changes, and assessing results from hydrological models (Potter, 1978). For these reasons, in continents like Europe, works were conducted to preserve and analyze such data (e.g., Bayliss et al., 2004; Pons et al., 2016). In South America, although hydrometric data have been taken in some rivers since the late nineteenth century, no or little effort was made to preserve these historical data which are mostly stored only in paper format. As this format restricts data access, only post-1900 hydrometric data were commonly used in studies on South American hydroclimate variability (e.g., Amarasekera et al., 1997; Antico et al., 2014, 2015; Berri et al., 2002; Camilloni & Barros, 2000; Dai et al., 2009; Dettinger & Diaz, 2000; Dettinger et al., 2001; García & Vargas, 1998; Labat et al., 2005; Pasquini & Depetris, 2007). Moreover, and most importantly, given the fragility of the data that are archived in paper format, there is a high risk in South America of losing unused relevant historical hydrometric data. Therefore, our objective here is to contribute to eliminate this risk by preserving the longest (last 143 years) hydrometric data set of the Paraná River basin, southeastern South America (see basin location in Figure 1).

To achieve our goal, we conducted the following rescue of paper format data of daily Paraná water levels observed at Rosario City, Argentina, from 1875 to present (see city location in Figure 1 and a nineteenth-century Rosario port photo in Figure 2). These data were first imaged and digitized to preserve the exact water level values reported in the paper documents. Next, the digitized level data were subjected to quality checks and homogeneity tests to identify wrong data values caused by observational and archiving problems. After this stage, we considered a correction for water level errors caused by gauge sinkings that may have occurred before 1909.

In addition to the water level data rescue described above, 41 discharge measurements taken at Rosario were obtained from previous works and they were used to estimate a rating curve (i.e., to estimate an empirical relationship between water level and discharge). This curve was used to convert water levels into

© 2018. American Geophysical Union. All Rights Reserved.

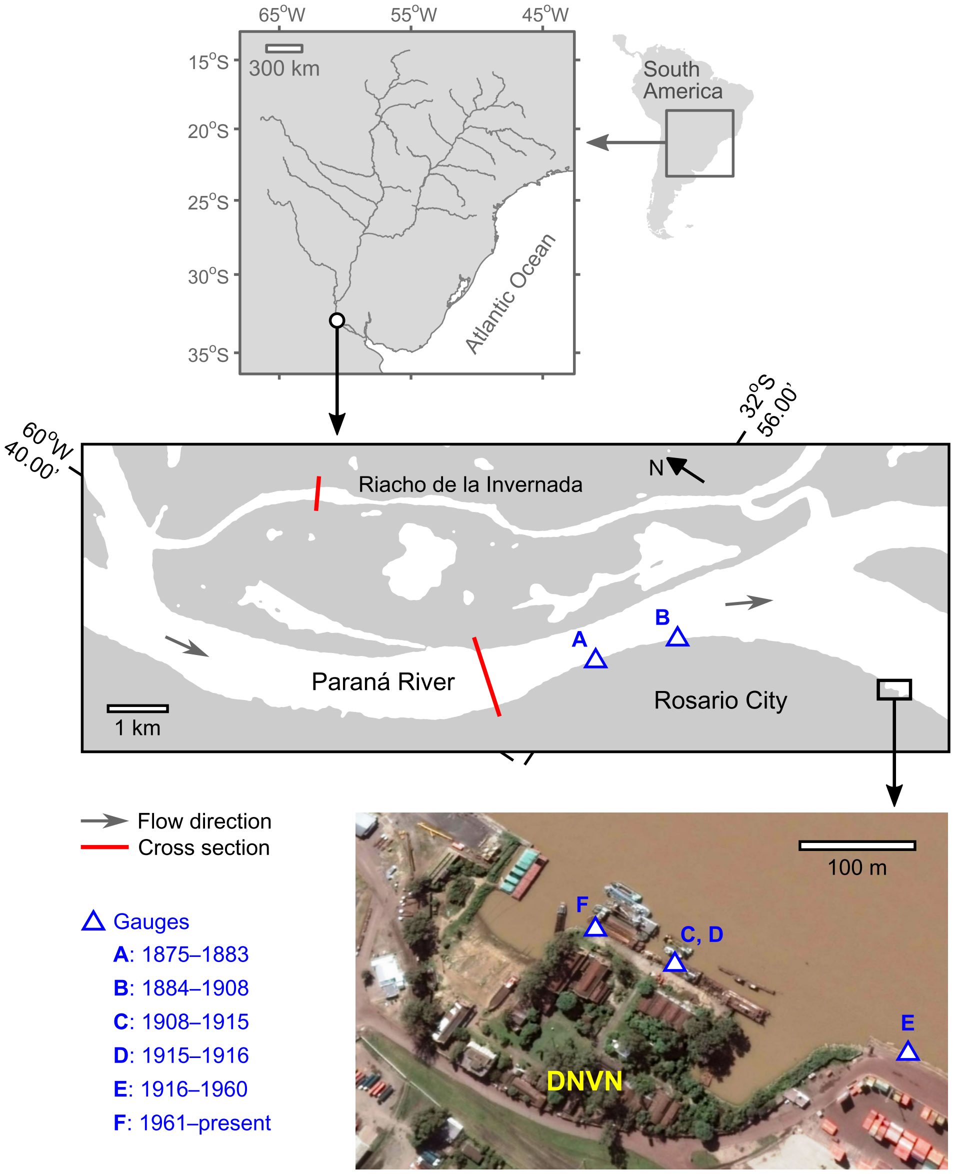


Figure 1. Location map of the Rosario gauges. (top) Drainage system of the Paraná River and location of Rosario City in this system (location shown by open circle). (middle) Paraná River at and near Rosario as depicted by a satellite Landsat-5 TM image (land/water boundary) acquired on 23 January 2011 (image downloaded from http://eros.usgs.gov/). (bottom) Satellite image of the DNVN facilities located in Rosario (image from Google Earth). The locations of Rosario gauges are indicated by triangles in the middle and bottom images. In the middle map, the names of streams (main and secondary) are indicated, and the red lines depict the cross sections that were usually used after 1920 for discharge measurements. The time spans for different gauges are indicated (complete gauge descriptions are in section 2 and Table 1).

discharges. It is noticed that river discharge data are relevant for climate and hydrological studies (see Hartmann et al., 2013, and references therein).

Since Rosario is relatively close to the Paraná River mouth (see Figure 1), gauges at this city register runoff integrated over most of the river basin. Considering this and the vast area of the Paraná basin (2.6×10^6 km^2, i.e., 60% of the European Union area), it is thus clear that Rosario hydrometric data provide valuable information on hydroclimate variability at subcontinental scales. Furthermore, the Paraná discharge is the third largest in South America and the sixth largest in the world (Schumm & Winkley, 1994), and hence

Figure 2. Rosario port photo from an album published in 1895 (Ferrazini & Co., 1895; available at Dr. Julio Marc Museum, Rosario).

Rosario data are also relevant for continental and global studies of the water cycle. These facts show the importance of the Paraná hydrometric data that were rescued in this work.

Interestingly, annual and monthly means of Paraná levels at Rosario for 1875–1915 and 1884–1916, respectively, are available in the printed tables presented by Mossman (1923). However, it should be noticed that this pioneering work only offered short segments of the 1875 to present Rosario level record. Moreover, Mossman (1923) did not provide daily water levels, which are important for studying extreme phenomena like floods. We also note that, during the writing of the present article, we did not find Rosario data in the following well-known repositories of hydrological data: (i) CLARIS LPB database (http://wp32.at.fcen.uba.ar/), (ii) data set of Dai et al. (2009) (http://www.cgd.ucar.edu/cas/catalog/surface/dai-runoff/), (iii) Global River Discharge Database (http://www.rivdis.sr.unh.edu/), (iv) Global Runoff Data Centre (http://www.bafg.de/GRDC/), and (v) Subsecretaría de Recursos Hídricos of Argentina (http://www.hidricosargentina.gov.ar/). Paraná levels at Rosario are provided by the National Water Agency of Brazil (http://www2.ana.gov.br/) but only for the short interval 1985–2003. We remedy these lacks of data by providing our rescued 1875 to present records of Paraná level and discharge at Rosario; these records are available at the Pangaea database (Antico et al., 2017).

In the following sections, we provide historical and technical information on Rosario gauges (section 2), describe how we rescued water level data from these gauges (section 3), explain the estimation of a rating curve (section 4), present the rescued hydrometric records (section 5), show applications of the rescued data to hydroclimatic studies (section 6), and give some concluding remarks (section 7).

2. Rosario Gauges

The port of Rosario became one of the most important commercial ports of Argentina in the last decades of the nineteenth century (Aguiar et al., 1995) (see Figure 2). Hence, not surprisingly, daily water levels started to be routinely recorded at this port in January 1875. From this time to present, observers have always been observing water levels at staff gauges. Table 1 describes the Rosario gauges whose data are presented in this work. The locations of these gauges are shown in Figure 1. The gauge A (1875–1883) was owned and operated by the FCCA, a British private railway company. Water levels from this gauge were originally observed and archived in British feet and they were subsequently converted to meters by Soulages and Duclout (1900). Since 1884, all the gauges considered here have been owned and operated by Argentinian national agencies, and levels have been always observed and archived in centimeters. For instance, the

Table 1
Description of the Rosario Gauges Considered in This Work

Gauge	Location[a]	Time span[b]	Spanish name[c]
A	60°38.613′W, 32°56.018′S	1 January 1875 to 31 December 1883	FCCA
B	60°37.929′W, 32°56.527′S	1 January 1884 to December 1908	Aduana/Subprefectura
C	60°37.097′W, 32°58.432′S	December 1908 to 1 January 1915	Escala de Estudios
D	60°37.097′W, 32°58.432′S	2 January 1915 to 13 November 1916	Hidrógrafo
E	60°37.057′W, 32°58.559′S	14 November 1916 to 31 December 1960	1.^{er} Muelle de Inflamables
F	60°37.098′W, 32°58.380′S	1 January 1961 to present	DNVN

[a] Obtained from Google Earth. All the coordinates are exact, except those for gauges C and D. The estimated location error for these two gauges is about ±80 m. [b] All the dates are exact, except the end date of gauge B, which is also the start date of gauge C. The exact day of this date is not known. [c] Spanish acronyms: FCCA, Central Argentine Railway; DNVN, national waterway agency.

Dirección Nacional de Vías Navegables (DNVN), which is the Argentinian national waterway agency, has been using the gauge F since 1961. The staff gauges A, B, C, D, and E were attached to wooden piers like that shown in Figure 2, and the most recent staff gauge F has been always attached to a reinforced concrete wall. Since 1926, the water levels have been observed every day at 6:00 AM local time at gauges E and F (S. Corsi, personal communication, 1986). For years before 1926, we could not find information about the specific hour of level observation.

It is noticed that the astronomical tides of La Plata Estuary (300 km downstream of Rosario) have no effect on the Paraná water level at Rosario (INCYTH, 1988; Révy, 1874). Conversely, a Rosario hydrometric data inspection (not shown here) revealed that some storm surges of La Plata Estuary (known as *sudestadas*) modify the water level at Rosario. However, these surges are short-lived (few days) and infrequent events (on average one moderate-to-strong event per year; Escobar et al., 2004). Therefore, most of time water level changes at Rosario gauges only reflect fluvial runoff changes; i.e., only reflect hydroclimate changes over the Paraná basin. For this reason, as mentioned in section 1 and shown in section 6, the Rosario hydrometric data rescued here are valuable for hydroclimate research.

3. Rescue of Water Level Data

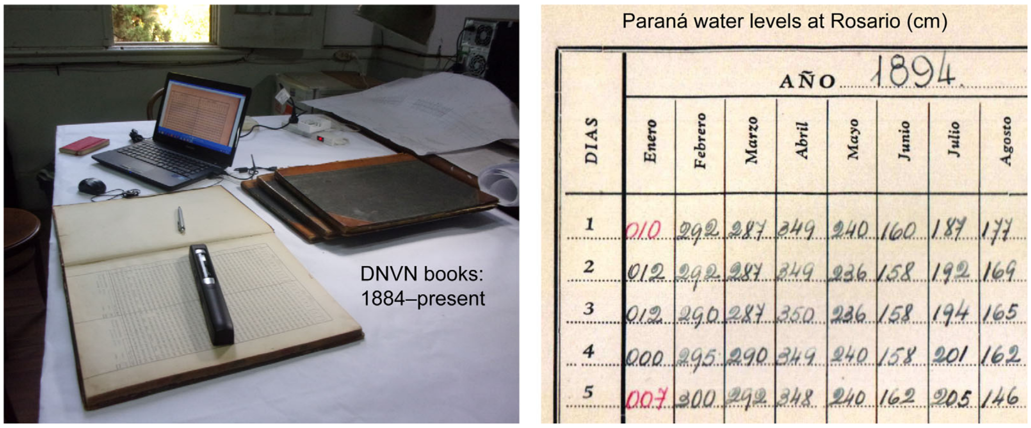
In this section, we describe our sources of paper format data of water level, and we explain how we imaged and digitized the data from these sources. We also present and discuss the results from applying quality checks and homogeneity tests to these data. At the end of this section, we present the correction that was proposed in a previous work to eliminate the water level errors caused by the possible sinkings of gauges A and B.

3.1. Sources of Official Water Level Data

All the Paraná water level observations rescued in this work were obtained from the documents that are described below in this section. Because these documents were produced by the Argentinian government, their water level data are of "official" character.

Our primary source of official level data consists of daily 1884 to present observations of Paraná level at Rosario that are handwritten in printed sheets of books held in the DNVN facilities in Rosario (see the DNVN books and one of their sheets in Figure 3). Every book page contains two sheets for two consecutive years. These sheets have a format that did not change through time (see the sheet format in right of Figure 3). In these sheets, water levels are handwritten as three-digit integer numbers that express levels in centimeters, and while positive values are in black or blue, negative levels are in red without the minus sign (see right of Figure 3); there are no missing values. According to the DNVN staff, the oldest sheets were filled sometime in the early twentieth century (the exact date is not known). The water levels written in DNVN books were observed at gauges B, C, D, E, and F, and they are all referred to the Rosario datum that is currently used by DNVN, hereafter mentioned as the DNVN datum.

According to old documents (e.g., Gould, 1884; Wolff, 1916), some Argentinian national agencies had printed sheets containing Paraná levels observed at gauge A before 1884. However, despite of our exhaustive search for these sheets, we could not find them. On the other hand, we were fortunate to find a governmental



Paraná water levels at Rosario (cm)

AÑO 1894.

DIAS	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto
1	010	292	287	349	240	160	187	177
2	012	292	287	349	236	158	192	169
3	012	290	287	350	236	158	194	165
4	000	295	290	349	240	158	201	162
5	007	300	292	348	240	162	205	146

Figure 3. DNVN books containing official data of daily Paraná levels observed at Rosario from January 1884 to present. (left) Imaging of book pages conducted in the DNVN facilities in Rosario and achieved with the portable scanner that is seen above the open book in the photo. (right) Scanned book page containing the daily level data sheet for the year 1894; in this sheet, rows are for days, columns are for months, handwritten levels are expressed in centimeters, and positive (negative) levels are in black (red).

document that depicts a figure of the time series of daily Paraná levels observed at gauge A from 1875 to 1883 and at gauge B from 1884 to 1899 (document B-VIIIa of Soulages and Duclout [1900], see this document in our Figure 4). The datum used to show 1875–1899 levels in this document is 1 m below the datum of DNVN-book levels (1884 to present), i.e., 1 m below the DNVN datum. This is stated in Repossini (1922a, 1922b) and confirmed here by a least squares linear regression over the overlapping data interval 1884–1899 (the regression was performed using the digitized levels obtained as described in next section). In this work, we refer all the 1875 to present Paraná levels to the DNVN datum.

3.2. Imaging and Digitization of Official Water Level Data

Original DNVN data sheets for 1884 to present were imaged with a portable scanner Nisuta NS-SCPO2 (see scanner in left of Figure 3). Color images were obtained with a resolution of 300 dpi. Figure 3 (right) shows an example of a scanned data sheet.

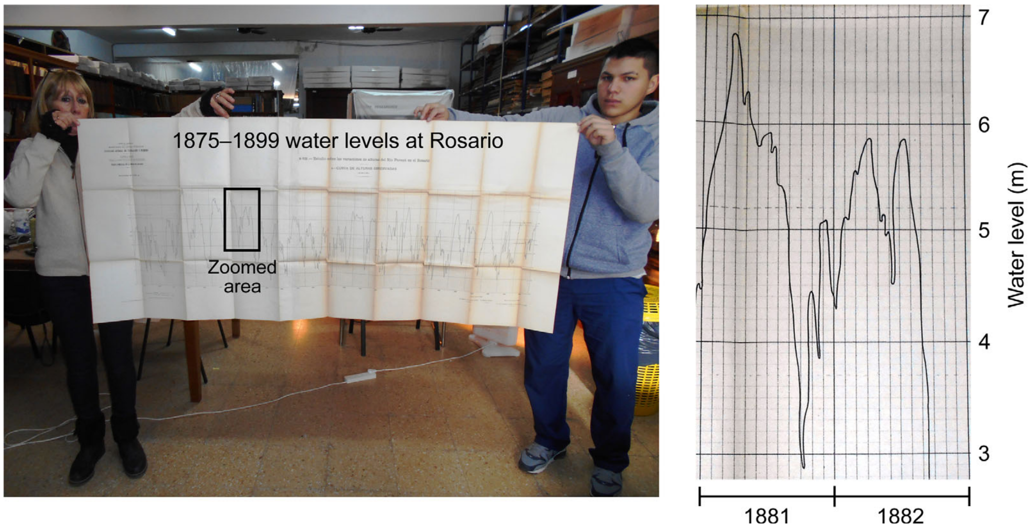


Figure 4. Document B-VIIIa of Soulages and Duclout (1900) depicting a figure of the time series of daily official Paraná levels observed at Rosario from January 1875 to December 1899. (left) Overall photo of this document taken at the Dr. Julio Marc Museum, Rosario (persons from left to right: Rita Bersisa and Sebastian Espindola). (right) Zoomed document area that shows daily official levels observed in 1881 and 1882; these levels should be reduced by one meter to refer them to the datum considered in this work (see section 3.1).

For handwritten observations like those considered here, key entry is recommended for digitizing data because it generates less errors than automated techniques such as speech and optical character recognition (Brönnimann et al., 2006). Hence, key entry was used here to digitize the handwritten daily official levels found in the DNVN books (1884 to present). To correct digitizing errors, one person digitized (key entered) the water level data and another person double checked each number; this method is recommended by Brönnimann et al. (2006). It is possible that few minor typing errors still remain in our 1884 to present record of daily official water levels. In this regard, we note that storing electronic images of the original data sheets (obtained as explained above) guarantee the possibility of future error corrections, even if the DNVN books are lost.

In addition to the data digitization described above, daily official Paraná levels at Rosario for 1875–1899 were directly digitized from the figure of Soulages and Duclout (1900) that is mentioned in section 3.1. The levels from this figure are reduced by 1 m to refer them to the DNVN datum (see section 3.1), and they are considered here only for the interval 1 January 1875 to 31 December 1883. For subsequent dates, we use the digitized DNVN-book level data, which are referred to the DNVN datum (see section 3.1).

3.3. Data Quality Checks

We applied four quality checks to the digitized data of daily official Paraná water level at Rosario. These checks and their results are described below in this section. These results are presented for all the years from 1875 to present, unless stated otherwise.

3.3.1. Gross Error Check

In the DNVN books, daily water levels in centimeters are handwritten as three-digit integer numbers (see section 3.1). Hence, we checked for the presence of suspicious integer numbers with less or more than three digits. We did not find any such suspicious case. The check described here was applied only over the interval 1884 to present because this is the time interval spanned by the DNVN books.

3.3.2. Range Check

Daily water levels at Rosario are expected to fluctuate between -2 and 7 m. All the digitized daily official levels lie within this range and, therefore, passed this check.

3.3.3. Step Check

A step is defined as the water level difference between two consecutive days. Since 1 January 1875, the absolute values of steps have always been less than 0.8 m, which is a reasonable limit for Rosario. Thus, we did not detect suspicious steps.

3.3.4. Flatliner Check

A flatliner is defined as a sequence of consecutive days with the same value. We flagged the values of suspicious flatliners in the following manner. If there is a flatliner with a duration of seven or more days, we flag all the flatliner values as suspicious except the first value that is kept unflagged. By doing this, we only flagged 73 daily values that correspond to less than 1% of the 1875 to present water level data. The maximum flatliner duration is of 9 days and there are only two flatliners of this length.

3.4. Data Homogeneity

Because different Rosario gauges were used during different periods (see Table 1), it is possible that some changes in the mean (break points) were introduced by observational problems such as an error committed in relating different datums. In this section, we investigate the presence of break points in the water level data, and we discuss whether or not the detected break points (if any) are caused by observational problems.

To detect break points, we applied the following homogeneity tests to the 1875–2016 record of annual means of official level (calculated using only the quality checked daily levels; see section 3.3): (i) Buishand test, (ii) Pettitt test (nonparametric rank test), and (iii) Standard normal homogeneity test (likelihood ratio test) (for test descriptions see Aguilar et al. [2003] and references therein). To implement these tests, we considered a significance level of 5% and 10,000 Monte Carlo simulations. The three tests detected a break point in 1972, when water level increased abruptly. After this detection, we split the time series of annual mean levels into two records (1875–1972 and 1973–2016), each of which was subjected to the three tests considered here. By doing so, only one more break point was detected in 1998 by two tests (Buishand and Pettitt). This break point corresponds to a water level reduction.

None of the detected break points (1972 or 1998) occurred at or near the times of the gauge changes presented in Table 1. We also applied the homogeneity tests considered above to a 1905–2015 record of

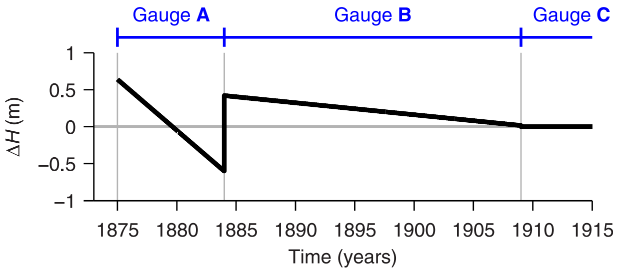


Figure 5. Water level correction (ΔH) proposed by Soulages (1948) to eliminate the water level errors caused by the possible sinkings of gauges A (1875–1883) and B (1884–1908). For a particular day of 1875–1908, ΔH should be added to the official water level observed at gauge A or B. After 31 December 1908, no correction is applied (i.e., $\Delta H=0$).

annual mean Paraná discharge at Paraná City, which is 160 km upstream of Rosario (discharge data were provided by the Subsecretaría de Recursos Hídricos from its website at http://www.hidricosargentina.gov.ar/). Once again, these tests detected the 1972 and 1998 break points described above; i.e., the mean of discharge (or equivalently of water level) increased in 1972 and decreased in 1998. Furthermore, Amsler et al. (2005) showed that the Paraná morphological changes that occurred in the 1970s are consistent with an increment of discharge and water level. All these results suggest that the 1972 and 1998 break points are not caused by instrumental problems. Indeed, these two break points are consistent with an observed relation between the Paraná runoff and the Interdecadal Pacific Oscillation (IPO) (Antico et al., 2014; Dettinger & Diaz, 2000; Dettinger et al., 2001; see also our section 6). According to this relation, the Paraná runoff decreased in the late 1990s due to an IPO transition from an El Niño-like IPO state to a La Niña-like IPO state. Conversely, in the 1970s, a sharp runoff increment followed an IPO transition from a La Niña-like IPO state to an El Niño-like IPO state. Besides the IPO-Paraná River link, other noninstrumental factors could have contributed to cause the 1972 break point. For instance, Krepper et al. (2008) suggested that the land use changes that occurred in the upper Paraná basin in the 1970s contributed to increase the Paraná runoff. In addition, several dams constructed in the Paraná basin in the 1970s could have also contributed to increase the annual mean discharge by preventing very low water levels (Anderson et al., 1993); during dry periods, dams can release water from their reservoirs.

From the results presented above, it is concluded that the water level break points detected in 1972 and 1998 do not have an instrumental cause. Therefore, we did not make any attempt to eliminate them from our data.

3.5. Correction for Possible Gauge Sinkings

According to the observational evidences presented by Soulages (1921, 1948) (e.g., results from leveling surveys), gradual sinkings were experienced by the wooden piers where gauges A and B were attached. To correct the errors introduced by these gauge movements into the water level record, Soulages (1948) estimated the correction that should be applied to the official water levels observed at gauges A and B. This correction is depicted in Figure 5. As stated in Repossini (1922b), the Argentinian government rejected the evidences and correction presented by Soulages (1921, 1948). Consequently, the official water level data described in section 3.1 do not include any correction for gauge sinkings. For the sake of completeness, in section 5 we present not only the official data but also a version of these data that is corrected as in Soulages (1948) to remove the errors caused by the possible sinkings of gauges A and B. In section 6, we show that these errors do not constitute a problem for the application of Rosario data in hydroclimate studies.

4. Rating Curve Estimation

A rating curve describes the empirical relationship between water level and discharge and, therefore, it can be used to convert water level observations into discharges. To estimate this curve, discharge measurements taken at different water levels are thus needed. In this section, we present (i) the discharge measurements that we obtained for Rosario from previous works, (ii) the way we used them to estimate a rating curve, and (iii) the uncertainties and errors of the discharges that were calculated using this curve.

At Rosario, the Paraná discharge measurements were usually obtained as the sum of discharges measured approximately simultaneously at two cross sections, one in the mainstream and the other one in a secondary stream (e.g., Riacho de la Invernada); see locations of the most recently used cross sections in Figure 1. Here we only consider discharge measurements taken when the official water level was five meters or less. This is to avoid situations in which large water volumes flowing over a 56 km wide floodplain are not captured by the river channel cross sections; these extreme floods are infrequent as they only represent 3% of the 1875 to present official water level data. Taking this into account, we were able to obtain 41 Rosario discharge measurements form previous works (Table 2 describes these measurements and provides the corresponding references). To perform all these measurements, current meters were always operated through the entire water column along every cross section. The current meter technology has evolved substantially

Table 2
Description of Paraná Discharge Measurements that Are Available for Rosario From Previous Works[a]

Work	*t*	*N*	*Q* (10^3 m^{3}s^{-1})	*H* (m)	Type of current meter
Duclout (1900)	1899	3	12.3–20.7	2.0 to 4.0	Mechanical
Repossini (1922b)	1904–1920	15	8.5–22.5	−0.3 to 5.0	Mechanical
Blanch et al. (1983)	1924–1969	22	4.7–25.4	−1.1 to 4.8	Mechanical
FICH-UNL (2006)	2 March 2006	1	13.7	2.2	ADCP

Note. ADCP, Acoustic Doppler Current Profiler.

[a] The following information about these data are presented: year(s) or date of measurements (*t*), number of measurements (*N*), discharge value or discharge range (*Q*), official water levels corresponding to discharge data (*H*), and type of current meter used for discharge measurements.

since the end of the nineteenth century and this probably caused a decrease with time of the errors of velocity and discharge measurements. We could not investigate this probable error reduction because we did not find the technical specifications of old current meters. Since all the discharge measurement considered here were taken by national agencies, they are referred to official water levels and thus they do not take into account the water level corrections presented in section 3.5. For the rating curve estimation presented below, we do not consider these corrections because they are small (less than 17 cm) after 1899, the year of the oldest discharge measurements, and are equal to zero after 1908.

To obtain a rating curve, a polynomial or power law function is usually fitted to the scatter plot of simultaneous observations of water level and discharge. According to the norm ISO 1100-2 (2010), an acceptable fit is obtained if (i) 15 or more data points are available for most of the observed range of water level variations, and if (ii) the level-discharge relation is not modified by river morphological changes during the data acquisition period. The first condition is satisfied here because, as it is shown in Table 2, there are 41 flow observations that encompass most of the observed range of water level fluctuations at Rosario (this range is −1.35 to 6.44 m). Immediately below, we assess whether the second condition is also satisfied by the available Rosario discharge measurements.

Since 1899, the first year of discharge data taking, the most important morphological change of the Paraná River at and near Rosario occurred in the late 1910s (1915–1920); this change is described in Appendix A. Thus, it is relevant to examine if this morphological change modified the water level-discharge relation because, if it did, different rating curves must be considered for different periods. To do this examination we compared the rating curves obtained for the intervals 1904–1920 and 1924–2006 because, while the first one corresponds to the hydraulic conditions that prevailed mainly before the morphological change, the second one corresponds to postchange conditions (see both curves in top and middle of Figure B1). As it is shown in bottom of Figure B1 and discussed in Appendix B, these two curves are not statistically significantly different from one another. That is, despite of the occurrence of a large morphological change, the water level-discharge relation did not change significantly and thus the condition (ii) mentioned in previous paragraph is satisfied.

Considering the above facts and results, we used all the Rosario discharge data described in Table 2 to obtain a single quadratic rating curve for the interval 1899–2006 (see top of Figure 6). The 95% confidence interval (95% CI) of this curve was estimated through a bootstrap method with replacement (1,000 replications) and is shown in top of Figure 6. For each water level value, the uncertainty of the rating curve was estimated by dividing the difference between the upper and lower limits of the 95% CI by the discharge calculated through the rating

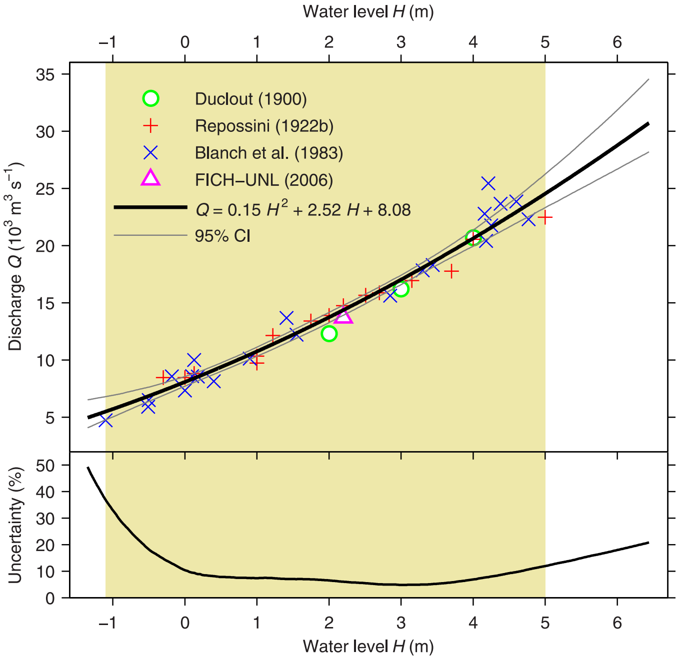


Figure 6. Rating curve at Rosario. (top) Quadratic rating curve obtained through a least squares fit to the discharge measurements shown in this figure and described in Table 2; the 95% confidence interval (95% CI) is shown and it was estimated as explained in section 4. (bottom) Discharge uncertainties estimated using the rating curve and its 95% CI (see section 4). In top and bottom, official water levels are considered, all the curves are depicted for the full water level range observed at Rosario (–1.35 to 6.44 m), and the shadow area indicates the levels encompassed by the discharge measurements (–1.1 to 5 m).

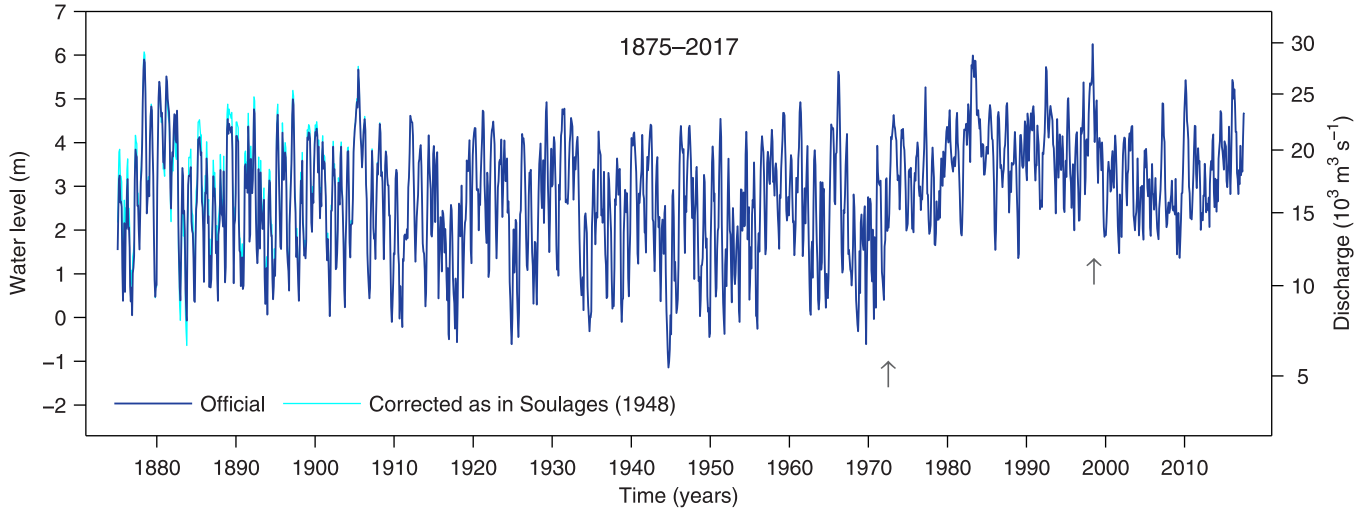


Figure 7. Monthly mean records of Paraná water level and discharge at Rosario for the interval January 1875 to July 2017. Official and corrected records are shown. The corrected water level record was obtained by applying the correction shown in Figure 5 to the official water level record (see section 3.5). Discharges were calculated from water levels (official and corrected) by using the rating curve depicted in top of Figure 6 (see sections 4 and 5). The arrows indicate the 1972 and 1998 break points described and discussed in section 3.4.

curve. As it can be observed in bottom of Figure 6, the uncertainty is less than 25% for levels greater than −0.76 m. In contrast, for lower levels, the uncertainty is relatively large (25–49%). It is noticed however that these lower levels are infrequent since they encompass less than 1% of the 1875 to present interval.

As it is mentioned in next section, the rating curve obtained here was used outside its temporal range to estimate discharges for years before 1899. A possible source of error in determining these discharges may be a rating curve modification (with respect to the curve of Figure 6) that could have happened in 1875–1898 if the river morphology had notably changed in this period. The Paraná River maps of 1884 and 1902 presented by Soulages (1944) show that no major river morphological changes occurred at and near Rosario from the 1880s to the early 1900s. Thus, it seems reasonable to use the 1899–2006 rating curve for the interval 1875–1898.

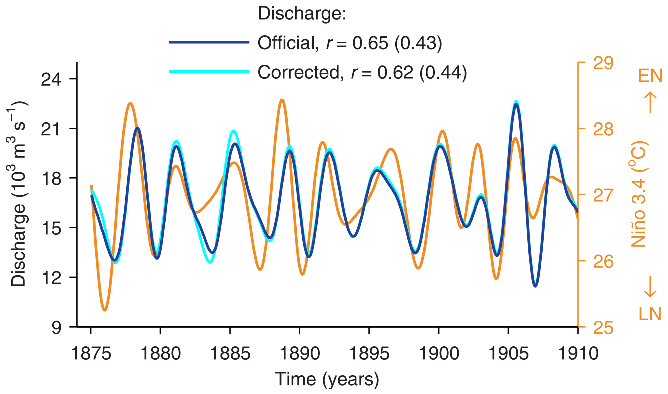


Figure 8. Band-pass filtered 1875–1909 records of Paraná discharge and ENSO Niño 3.4 index (Trenberth & Stepaniak, 2001). A Fourier band-pass filter with cutoff periods of 2 and 7 years was applied to monthly mean records of Niño 3.4 index and discharge (official and corrected) in order to visualize oscillations with ENSO frequencies. El Niño (EN) and La Niña (LN) ENSO events are indicated. For each discharge record, its zero-lag Pearson correlation coefficient (r) with the Niño 3.4 series is shown along with the 95% confidence threshold in parentheses. This threshold is estimated by combining 1,000 Monte Carlo iterations with frequency-domain time series modeling so that autocorrelation is considered (Macias-Fauria et al., 2012). If r is greater than its 95% confidence threshold, then r is statistically different from zero.

5. The Rescued Hydrometric Records

The major outcome of the data rescue described in section 3 is a record of official daily Paraná levels observed at Rosario from 1 January 1875 to 31 July 2017. A corrected version of these official levels was obtained as described in section 3.5 by removing the errors caused by the possible sinkings of gauges A (1875–1883) and B (1884–1908). The rating curve presented in section 4 was applied to the official and corrected daily water levels to estimate daily discharges; for brevity, hereafter we use the term “official discharges” for the first case, and the term “corrected discharges” for the second case. For each variable, monthly means were calculated by using only the daily values that passed all the quality checks described in section 3.3; for every month of the observational interval (January 1875 to July 2017), there are at least 22 daily values that satisfy this condition. Figure 7 shows the obtained monthly mean records of official and corrected water levels and discharges. In this figure, the 1972 and 1998 break points described in section 3.4 are clearly discernible.

6. Examples of Applications to Hydroclimate Studies

In this section, we present examples of useful applications of the rescued data to hydroclimate studies. As shown next, these examples

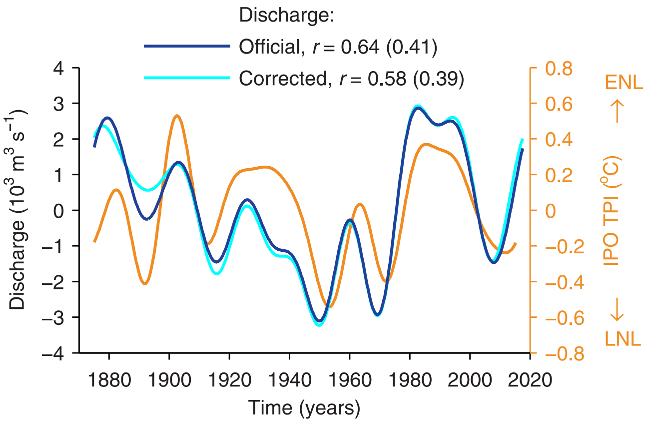


Figure 9. Low-pass filtered 1875–2015 records of Paraná discharge and IPO tripole index (IPO TPI) (Henley et al., 2015). A Fourier low-pass filter with a cutoff period of 15 years was applied to linearly detrended monthly mean records of IPO TPI and discharge (official and corrected). El Niño-like (ENL) and La Niña-like (LNL) states of IPO are indicated. For each discharge record, its zero-lag Pearson correlation coefficient (*r*) with the IPO TPI series is shown along with the 95% confidence threshold in parentheses. This threshold is calculated as explained in caption of Figure 6.

consist of analyses of our Paraná discharge data that reveal new insights into the Paraná-Pacific Ocean links reported in previous studies (see references below in this section).

Many former studies for the 20th and 21st centuries showed that the interannual Pacific climate cycle known as El Niño/Southern Oscillation (ENSO; oscillatory period of 2–7 years) has an influence on the Paraná discharge (e.g., Amarasekera et al., 1997; Antico et al., 2014; Berri et al., 2002; Camilloni & Barros, 2000; Dai et al., 2009; Dettinger & Diaz, 2000). Discharge is high during El Niño ENSO events, and low during La Niña ENSO events; physical mechanisms underlying this association are described in Garreaud et al. (2009). Our data analysis shown in Figure 8 reveals that this ENSO-discharge link existed not only after 1900 but also during the last 25 years of the nineteenth century.

At interdecadal time scales, and for post-1900 years, previous works detected a relation between the Paraná discharge and the IPO, a slow ENSO-like climate cycle that is also known as the Pacific Decadal Oscillation (Antico et al., 2014; Dettinger & Diaz, 2000; Dettinger et al., 2001). The Paraná discharge is enhanced during El Niño-like IPO years, and diminished during La Niña-like IPO years. Our Paraná discharge records support this relation because, as revealed by the data analysis of Figure 9, these records and the IPO index series have similar interdecadal cycles. In particular, Figure 9 reveals that this similitude existed in last three decades of the nineteenth century. To the best of our knowledge, this is the first time that an evidence of an IPO-Paraná discharge link is reported for the nineteenth century. It can be also noticed in Figure 9 that the rise and fall of discharge observed in the 1970s and late 1990s, respectively, and their associated IPO changes support the interpretations given in section 3.4 for the 1972 and 1998 break points.

As it can be observed in Figures 8 and 9, the new insights presented above in this section are obtained either from the official or from the corrected discharge data. This shows that the possible observational errors presented in section 3.5 do not pose a problem for the use of our rescued data in hydroclimate research.

7. Concluding Remarks

We have conducted the first rescue of paper format data of daily Paraná River water levels observed at Rosario City (Argentina) from 1875 to present. This rescue was achieved not only by imaging and digitizing paper format data but also by considering errors committed during the acquisition and transcription of water level observations. Additionally, a rating curve was obtained for Rosario and it was used to convert water levels into discharges. The water level and discharge data rescued in this work are available at the Pangaea database (Antico et al., 2017).

In regard to the quality of our data, we notice that while daily water levels for 1884 to present were digitized from written numerical values, those for 1875–1883 were directly digitized from a figure made in 1900. For this reason, the water levels and discharges for 1875–1883 are expected to be less precise than those for subsequent years. Moreover, pre-1909 hydrometric data may include errors caused by gauge sinking (a correction for these errors is presented in section 3.5). It is also worth to stress that while the rescued water level observations are valuable for studying floods and extraordinary low level episodes, our discharge estimates for these extreme events should be used with caution because they have large uncertainties. Nevertheless, despite these drawbacks, we proved that our rescued data provide useful information for hydroclimate research. We did this by performing analyses of our discharge data that reveal new insights into the interannual and interdecadal influences of the Pacific Ocean on the Paraná discharge (see section 6).

This work offers the longest (last 143 years) continuous hydrometric records of the Paraná River and, as far as we are aware, no other large South American river has such long records. This opens the possibility of comparing these new Paraná records with hydrometric records of similar time span that are available for other important rivers of the world like the Danube (Europe), the Mississippi (North America), the Murray (Australia), and the Saint Lawrence (North America). A comparison of this kind would certainly contribute to hydroclimate research and thus could be the subject of future studies. With respect to the contribution of this work to paleoclimate studies, we mention that the long instrumental records rescued here will be valuable for future calibrations of proxies for Paraná flow changes. In addition to these potential contributions to hydroclimatology and paleoclimatology, our pre-1900 data would also be used to extend previous studies on the twentieth-century Paraná morphological response to interdecadal discharge changes (e.g., Amsler et al., 2005).

In order to facilitate a future combination of our discharge data with others, we used classical and widely used techniques to estimate discharges from water levels, and to calculate the associated discharge uncertainties. However, some more sophisticated methods were recently proposed to do these tasks (e.g., Coxon et al., 2015; Westerberg & McMillan, 2015). We leave for futures works the possible application of these new techniques to the Paraná hydrometric data presented here.

Finally, we mention that this work is part of a project that is rescuing historical hydrometric data not only in the Paraná basin but also in the Uruguay basin, another large river basin of South America (continuous Uruguay records start in 1892; see project information in https://www.idare-portal.org/). It is likely that similar historical data but for other South American rivers are still stored only in paper format in some archives. We then hope that our work will motivate future efforts dedicated to localize and rescue such data.

Appendix A: The Paraná Morphological Change of 1915–1920

The most important morphological change that was documented in the Paraná River at and near Rosario occurred mainly between 1915 and 1920. This change was caused by a combination of intense engineering works and natural processes. Here we briefly describe this morphological change. A more detailed description can be found in Repossini (1922a).

Top of Figure A1 shows that, at the end of the nineteenth century, the Espinillo Island was in front of Rosario City and La Invernada Island was immediately upstream of the Espinillo Island. The 51% of the total river discharge flowed between these two islands and though the channel that was east of the Espinillo Island. The remaining 49% of the discharge flowed between the Espinillo Island and the Rosario coast. The Riacho de la Invernada stream, east of La Invernada Island, was closed.

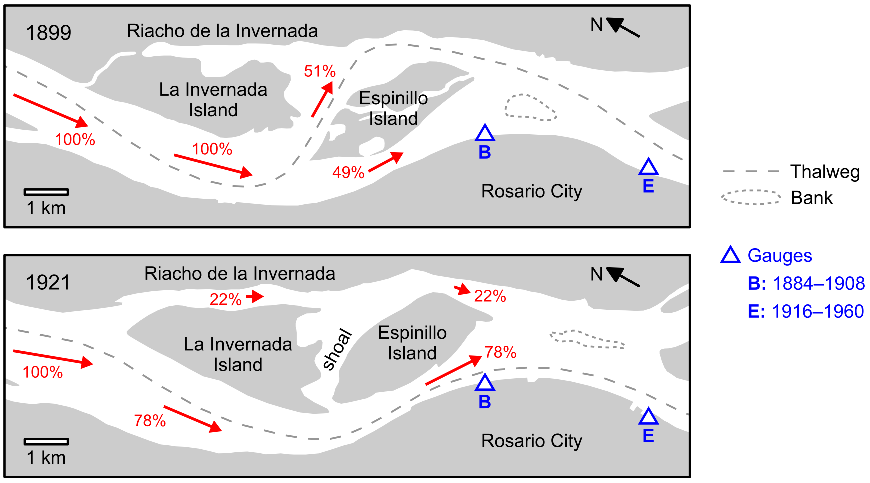


Figure A1. Paraná River maps of 1899 and 1921. (top) Map depicting the channels, islands, and banks that existed at and near Rosario in 1899. (bottom) Like top but for 1921. In top and bottom, the long-dashed lines depict the main thalweg, the triangles indicate the locations of gauges B and E, and the red numbers are the fractions of total discharge corresponding to different channels. The red arrows indicate the flow direction; arrow lengths are proportional to the fraction of total discharge. Both maps are adapted from Repossini (1922a).

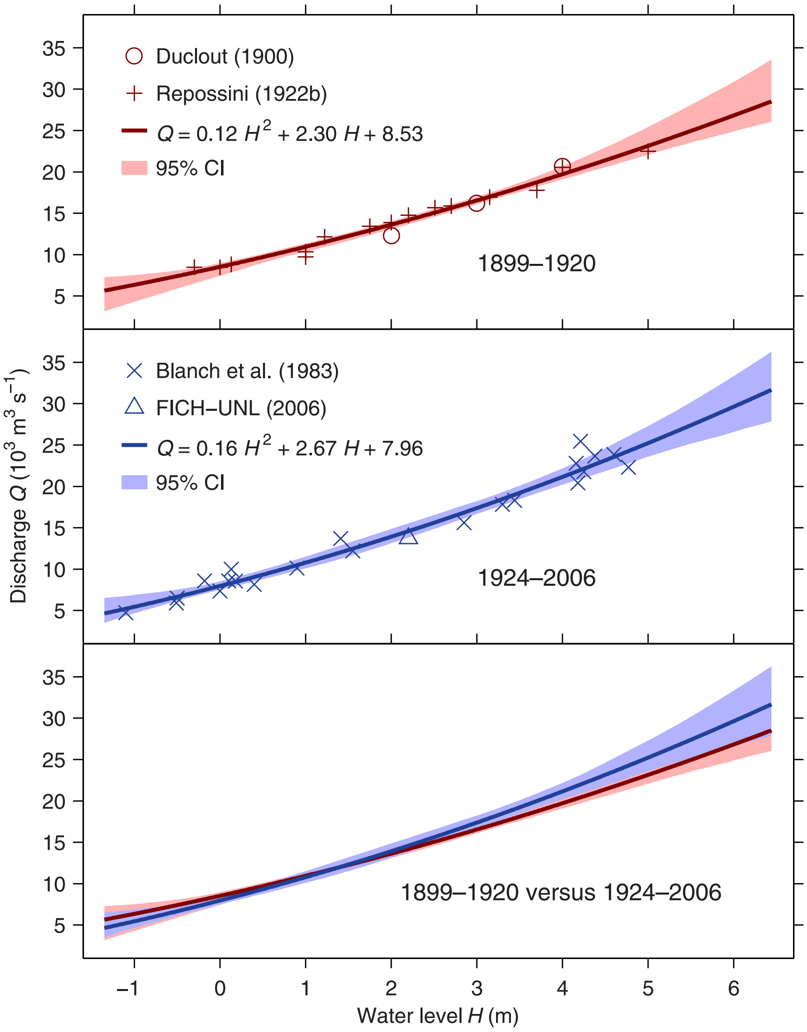


Figure B1. Rating curves at Rosario for 1899–1920 and 1924–2006. (top) Quadratic rating curve obtained through a least squares fit to the flow measurements taken at Rosario during the interval 1899–1920. (middle) Same as top but for the interval 1924–2006. (bottom) Comparison of the rating curves for 1899–1920 and 1924–2006. In the three figures, official water levels are considered, and all the curves are depicted for the full water level range observed at Rosario (–1.35 to 6.44 m). The discharge measurements shown in top and middle are described in Table 2. For each rating curve, the 95% confidence interval (95% CI) is shown as a shadow and it was estimated through a bootstrap method with replacement (1,000 replications).

From 1915 to 1920, the channel between the Espinillo and La Invernada Islands was closed by sedimentation processes and the Riacho de la Invernada stream was opened. As it is shown in bottom of Figure A1, in 1921 the two mentioned islands were connected by a shoal so that they formed one single large island. Most of the river discharge flowed between this large island and the Rosario coast. This hydraulic setting persisted until present.

Appendix B: Rating Curve Stability

In this appendix, we discuss whether the rating curve at Rosario was significantly modified by the 1915–1920 morphological change described in Appendix A or whether the rating curve remained approximately stable (i.e., unchanged).

The discharge measurements described in Table 2 allowed us to obtain two rating curves, one corresponding to the hydraulic conditions that prevailed mainly before the morphological change mentioned above (Figure B1, top), and another one for postchange conditions (Figure B1, middle). It is noticed that each of these curves was obtained by using more than 15 discharge measurements that encompass most of the observed range of water level fluctuations at Rosario. According to the norm ISO 1100-2 (2010), this implies that the available discharge data are suitable for an appropriate determination of the two rating curves that are shown in top and middle of Figure B1. Considering this, we present a comparison of these two curves in bottom of Figure B1. As it can be observed in this plot, the 95% confidence intervals of the two rating curves overlap at every water level, implying that the two curves are not statistically significantly different from one another. Therefore, it can be concluded that the rating curve at Rosario was not significantly modified by the major morphological change that occurred in the late 1910s at and near Rosario. That is, the rating curve remained approximately stable or unchanged.

Acknowledgments

We acknowledge the dedicated work of former and present DNVN-Rosario staff in collecting and archiving 1884 to present hydrometric data. We are also indebted to the following Argentinian individuals for helping us to find historical data and information: Juan M. Blanch (deceased), from DNVN-Rosario and UTN-Rosario; Sabino Corsi (deceased), DNVN-Rosario and FICH-UNL; Jorge E. Deniri, Corrientes Province Archive; Carolina Gregui, Argentinian Centre of Engineers; Carlos Mey, Histarmar Foundation; Miriam Mitri, Santa Fe Province Archive; Daniel Olmos, Argentinian National Weather Service; and Andrés R. Rousseaux, retired from Argentinian Coast Guard. We particularly thank Ester E. Dadidov and Rita Bersisa, both from Dr. Julio Marc Museum at Rosario, for facilitating the access to documents with 1875–1921 hydrometric data. Jazmin Vieytes is thanked for key entering data. El Niño 3.4 index data were provided by the GCOS Working Group on Surface Pressure from its website at http://www.esrl.noaa.gov/psd/gcos_wgsp/. The tripole IPO index record, derived from HadISST1 data, was provided by NOAA-ESRL from their website at http://www.esrl.noaa.gov/. The hydrometric data rescued in this work are available at https://doi.pangaea.de/10.1594/PANGAEA.882613. This work was partly financed by the Universidad Nacional del Litoral through the CAI + D and PIRHCa programs. Finally, but not least, we are grateful to the Editor (Jessica Lundquist), an Associate Editor, and two anonymous reviewers for their constructive comments that improved the manuscript.

References

Aguiar, R. O., Alfano, C. D., Estevez, D. M., Laferrara, O. D., & Seselovsky, E. R. (1995). Anexo V: hechos históricos específicos. In J. Alemany Llovera (Study Director), *Estudio para el desarrollo del puerto de Rosario* (pp. 308–325). Rosario, Argentina: Fundación Banco Municipal de Rosario.

Aguilar, E., Auer, I., Brunet, M., Peterson, T. C, & Wieringa, J. (2003). *Guidelines on climate metadata and homogenization* (WMO-TD No. 1186). Geneva, Switzerland: World Meteorological Organization.

Amarasekera, K. N., Lee, R. F., Williams, E. R., & Eltahir, E. A. B. (1997). ENSO and the natural variability in the flow of tropical rivers. *Journal of Hydrology*, *200*, 24–39.

Amsler, M. L., Ramonell, C. G., & Toniolo, H. A. (2005). Morphologic changes in the Paraná River channel (Argentina) in the light of the climate variability during the 20th century. *Geomorphology*, *70*, 257–278. https://doi.org/10.1016/j.geomorph.2005.02.008

Anderson, R. J., da Franca Ribeiro dos Santos, N., & Diaz, H. F. (1993). *An analysis of flooding in the Paraná/Paraguay River basin* (LATEN Dissemination Note No. 5). World Bank.

Antico, A., Aguiar, R. O., & Amsler, M. L. (2017). *Paraná River hydrometric data from Rosario City, Argentina: 1875 to 2017. PANGAEA*. https://doi.pangaea.de/10.1594/PANGAEA.882613

Antico, A., Schlotthauer, G., & Torres, M. E. (2014). Analysis of hydroclimatic variability and trends using a novel empirical mode decomposition: Application to the Paraná River Basin. *Journal of Geophysical Research: Atmosphere*, *119*, 1218–1233. https://doi.org/10.1002/2013JD020420

Antico, A., Torres, M. E., & Diaz, H. F. (2015). Contributions of different time scales to extreme Paraná floods. *Climate Dynamics*, *46*, 3785–3792. https://doi.org/10.1007/s00382-015-2804-x

Bayliss, A., Norris, J., & Marsh, T. (2004). The Wendover Springs record: An insight into the past and a benchmark for the future. *Weather*, *59*, 267–271. https://doi.org/10.1256/wea.56.04

Berri, G. J., Ghietto, M. A., & García, N. O. (2002). The influence of ENSO in the flows of the upper Paraná River of South America over the past 100 years. *Journal of Hydrometeorology*, *3*, 57–65.

Blanch, J. M., Pacilio, P. D., & Scioli, G. (1983). *Río Paraná - Curva de caudales frente puerto Rosario* (Document No. 5129 P.R.1). Argentina: Ministerio de Obras y Servicios Públicos, Dirección Nacional de Construcciones Portuarias y Vías Navegables, Departamento Paraná Inferior.

Brönnimann, S., Annis, J., Dann, W., Ewen, T., Grant, A. N., Griesser, T., . . . Vogler, C. (2006). A guide for digitising manuscript climate data. *Climate of the Past*, *2*, 137–144. https://doi.org/10.5194/cp-2-137-2006

Camilloni, I. A., & Barros, V. R. (2000). The Paraná River response to El Niño 1982–83 and 1997–98 events. *Journal of Hydrometeorology*, *1*, 412–430.

Coxon, G., Freer, J., Westerberg, I. K., Wagener, T., Woods, R., & Smith, P. J. (2015). A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations. *Water Resources Research*, *51*, 5531–5546. https://doi.org/10.1002/2014WR016532

Dai, A., Qian, T., Trenberth, K. E., & Milliman, J. D. (2009). Changes in continental freshwater discharge from 1948 to 2004. *Journal of Climate*, *22*, 2773–2792. https://doi.org/10.1175/2008JCLI2592.1

Dettinger, M. D., Battisti, D. S., Garreaud, R. D., McCabe, G. J., & Bitz, C. M. (2001). Interhemispheric effects of interannual and decadal ENSO-like climate variations on the Americas. In V. Markgraf (Ed.), *Interhemispheric climate linkages: Present and past climates in the Americas and their societal effects* (pp. 1–16). Cambridge, MA: Academic Press.

Dettinger, M. D., & Diaz, H. F. (2000). Global characteristics of stream flow seasonality and variability. *Journal of Hydrometeorology*, *1*, 289–310.

Duclout, J. (1900). Memoria sobre las pendientes y velocidades del Río Paraná frente al Rosario. In *Concurso para la construcción y explotación de un puerto comercial en la ciudad de Rosario*. Buenos Aires, Argentina: Ministerio de Obras Públicas, Inspección General de Navegación y Puertos, 26–118.

Escobar, G., Vargas, W., & Bischoff, S. (2004). Wind tides in the Rio de la Plata estuary: Meteorological conditions. *International Journal of Climatology*, *24*, 1159–1169. https://doi.org/10.1002/joc.1026

Ferrazini & Co. (1895). *Vistas de la ciudad de Rosario de Santa Fé* (2nd ed.). Rosario, Argentina: Ferrazini & Co.

FICH-UNL. (2006). *Estudios complementarios hidráulicos y morfológicos de la zona de la Isla de la Invernada*. Santa Fe, Argentina: FICH-UNL.

García, N. O., & Vargas, W. M. (1998). The temporal climatic variability in the 'Río De La Plata' basin displayed by the river discharges. *Climatic Change*, *38*, 359–379.

Garreaud, R. D., Vuille, M., Compagnucci, R., & Marengo, J. (2009). Present-day South American climate. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *281*, 180–195. https://doi.org/10.1016/j.palaeo.2007.10.032

Gould, B. A. (1884). *Anales de la oficina meteorológica Argentina, Tomo IV*. Buenos Aires, Argentina: Oficina Meteorológica Argentina.

Hartmann, D., Klein Tank, A., Rusticucci, M., Alexander, L., Brönnimann, S., Charabi, Y., . . . Zhai, P. (2013). Observations: Atmosphere and surface. In *Climate change 2013: The physical science basis. Contribution of Working Group I to the fifth assessment report of the Intergovernmental Panel on Climate Change, edited by T. Stocker, D. Qin, G. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley* (pp. 161–218). Cambridge, UK: Cambridge University Press.

Henley, B. J., Gergis, J., Karoly, D. J., Power, S., Kennedy, J., & Folland, C. K. (2015). A tripole index for the Interdecadal Pacific Oscillation. *Climate Dynamics*, *45*, 3077–3090. https://doi.org/10.1007/s00382-015-2525-1

INCYTH. (1988). *Estudio del Paso Las Hermanas, Segundo Informe de Avance*. Argentina: Laboratorio de Hidráulica Aplicada, Instituto Nacional de Ciencia y Técnica Hídricas.

ISO 1100-2. (2010). *Hydrometry—Measurement of liquid flow in open channels—Part 2: Determination of the stage-discharge relationship*. Geneva, Switzerland: International Organization for Standardization.

Krepper, C. M., García, N. O., & Jones, P. D. (2008). Low-frequency response of the upper Paraná basin. *International Journal of Climatology*, *28*, 351–360. https://doi.org/10.1002/joc.1535

Labat, D., Ronchail, J., & Guyot, J. L. (2005). Recent advances in wavelet analyses: Part 2—Amazon, Parana, Orinoco and Congo discharges time scale variability. *Journal of Hydrology*, *314*, 289–311. https://doi.org/10.1016/j.jhydrol.2005.04.004

Macias-Fauria, M., Grinsted, A., Helama, S., & Holopainen, J. (2012). Persistence matters: Estimation of the statistical significance of paleoclimatic reconstruction statistics from autocorrelated time series. *Dendrochronologia*, *30*, 179–187. https://doi.org/10.1016/j.dendro.2011.08.003

Mossman, R. C. (1923). *On Indian monsoon rainfall in relation to South American weather, 1875–1914* (Vol. 23, pp. 157–243). Simla, India: Superintendent Government Printing.

Pasquini, A. I., & Depetris, P. J. (2007). Discharge trends and flow dynamics of South American rivers draining the southern Atlantic seaboard: An overview. *Journal of Hydrology*, *333*, 385–399. https://doi.org/10.1016/j.jhydrol.2006.09.005

Pons, F., Laroche, C., Trmal, C., Puechberty, R., & Baillon, M. (2016). Hydrometry data rescue, a stake for the future, *E3S Web of Conferences*, *7*, 04021.

Potter, H. R. (1978). *The use of historic records for the augmentation of hydrological data* (Rep. 46). Wallingford, UK: Institute of Hydrology.

Repossini, J. (1922a). El río Paraná frente a Rosario - Su corrección presente y futura - Los derrumbamientos sobre el brazo de la Ciudad y sus causas - El porvenir del Puerto de Rosario. *La Ingeniería*, *571*, 272–281.

Repossini, J. (1922b). Los planos y documentos del M.O.P. relativos al río Paraná. *La Ingeniería*, *567*, 24–30.

Révy, J. J. (1874). *Hydraulics of great rivers: The Paraná, the Uruguay, and the La Plata Estuary*. London: E. & FN Spon.

Schumm, S. A., & Winkley, B. R. (1994). The character of large alluvial rivers. In S. A. Schumm & B. R. Winkley (Eds.), *The variability of large alluvial rivers* (pp. 1–9). New York: American Society of Civil Engineers.

Soulages, E. (1921). Un error técnico - Los planos y documentos del M.O.P. relativos al río Paraná – V. *La Ingeniería*, *564*, 441–447.

Soulages, E. (1944). El puerto de Rosario en peligro por la evolución natural de la configuración del cauce del río: Conclusión. *La Ingeniería*, *841*, 896–900.

Soulages, E. (1948). Las fluctuaciones del nivel del río Paraná: Los hundimientos de las escalas hidrométricas en el puerto de Rosario. *La Ingeniería*, *884*, 291–295.

Soulages, E., & Duclout, J. (1900). Memoria sobre las variaciones de altura del Río Paraná frente al Rosario. In *Concurso para la construcción y explotación de un puerto comercial en la ciudad de Rosario*. Buenos Aires, Argentina: Ministerio de Obras Públicas, Inspección General de Navegación y Puertos, 1–17.

Trenberth, K. E., & Stepaniak, D. P. (2001). Indices of El Niño evolution. *Journal of Climate*, *14*, 1697–1701.

Westerberg, I. K., & McMillan, H. K. (2015). Uncertainty in hydrological signatures. *Hydrology and Earth System Sciences*, *19*, 3951–3968. https://doi.org/10.5194/hess-19-3951-2015

Wolff, E. (1916). *Informe de la Oficina Hidrométrica* (Boletín Mensual No. 1). Buenos Aires, Argentina: Oficina Meteorológica Nacional.

Erratum

In the originally published version of this article, the fourth row of Table 2 incorrectly referred to “ACDP.” This has been corrected, and this may be considered the official version of record.