Biogeochemistry
https://doi.org/10.1007/s10533-019-00585-2

Decreased buffering capacity and increased recovery time for legacy phosphorus in a typical watershed in eastern China between 1960 and 2010

Dingjiang Chen · Yufu Zhang · Hong Shen · Mengya Yao · Minpeng Hu · Randy A. Dahlgren

Received: 23 February 2019 / Accepted: 31 July 2019

Abstract Legacy phosphorus (P) accumulated in watersheds from excessive historical P inputs is recognized as an important component of water pollution control and sustainable P management in watersheds worldwide. However, little is known about how watershed P buffering capacity responds to legacy P pressures over time and how long it takes for riverine P concentrations to recover to a target level, especially in developing countries. This study examined long-term (1960–2010) accumulated legacy P stock, P buffering capacity and riverine TP flux dynamics to predict riverine P-reduction recovery times in the Yongan watershed of eastern China. Due to a growing legacy P stock coupled with changes in land use and climate, estimated short- and long-term buffering metrics (i.e., watershed ability to retain current year and historically accumulated surplus P, respectively) decreased by 65% and 36%, respectively, resulting in a 15-fold increase of riverine P flux between 1980 and 2010. An empirical model incorporating accumulated legacy P stock and annual precipitation was developed ($R^2 = 0.99$) and used to estimate a critical legacy P stock of 22.2 ton P km^{-2} (95% CI 19.4–25.3 ton P km^{-2}) that would prevent exceedance of a target riverine TP concentration of 0.05 mg P L^{-1}. Using an exponential decay model, the recovery time for depleting the estimated legacy P stock in 2010 (29.3 ton P km^{-2}) to the critical level (22.2 ton P km^{-2}) via riverine flux was 456 years (95% CI 353–560 years), 159 years (95% CI 57–262 years) and 318 years (95% CI 238–400 years) under scenarios of a 4% reduction in annual P inputs, total cessation of P inputs, and 4% reduction of annual P inputs with a 10% increase in average annual precipitation, respectively. Given the lower P buffering capacity and lengthening recovery time, strategies to reduce P inputs and utilize soil legacy P for crop production are necessary to effectively control riverine P pollution and conserve global rock P resources. A long-term perspective that incorporates both contemporary and historical information is required for developing sustainable P management strategies to optimize both agronomic and environmental benefits at the watershed scale.

Responsible Editor: Amy M. Marcarelli

D. Chen (✉) · Y. Zhang · H. Shen · M. Yao · M. Hu
College of Environmental Science and Resources, Zhejiang University, Hangzhou 310058, China
e-mail: chendj@zju.edu.cn

D. Chen
Ministry of Education Key Laboratory of Environment Remediation and Ecological Health, Zhejiang University, Hangzhou 310058, China

M. Hu
Zhejiang Provincial Key Laboratory of Subtropical Soil and Plant Nutrition, Zhejiang University, Hangzhou 310058, China

R. A. Dahlgren
Department of Land, Air, and Water Resources, University of California, Davis, CA 95616, USA

Published online: 05 August 2019

Keywords Phosphorus · Legacy nutrients · Watershed · Phosphorus buffering capacity · Eutrophication · Lag time

Introduction

Phosphorus (P) is an essential element for all forms of life on earth and plays an important role in global agricultural production (Jarvie et al. 2014; Withers et al. 2014). Anthropogenic P addition, which is critical for producing food, fiber, and bioenergy to meet the needs of a rapidly increasing global population, has led to massive changes in the P cycle at both local and global scales (MacDonald et al. 2011; Powers et al. 2016). In many watersheds, anthropogenic P inputs have greatly exceeded P uptake by crop and animal products, which increases P transfer to surface waters where it may lead to eutrophication, hypoxia and coastal dead zones (Carpenter 2008; Conley et al. 2009; Vitousek et al. 2010; Howarth et al. 2011). Given rising food demands by an increasing population and shortages of rock P reserves, sustainable P management strategies for optimizing both agronomic and environmental benefits are a critical global concern (Chen et al. 2018).

Due to the inherent lithophilic properties of P (i.e., P is subject to sorption or coprecipitation), excess P has a propensity to accumulate in soils and sediments (Goyette et al. 2018; Chen et al. 2018). The accumulated P increases P transfers along the land–freshwater continuum, causing a considerable lag time between anthropogenic P inputs and P delivery to surface waters (Kleinman et al. 2011; Goyette et al. 2018). For example, global P accumulated in croplands was 815 Tg P between 1965 and 2007 (Sattari et al. 2012), which was $\sim$ 40-fold higher than applied P fertilizer in 2012 (20.3 Tg P year^{-1}). Such accumulated anthropogenic P over time, known as legacy P, can be remobilized or recycled by changes in land management and climate (Meals et al. 2010; Sharpley et al. 2013) or landscape P saturation after long-term additions of excessive P levels (Kleinman et al. 2011). Increasing evidence indicates that legacy nutrient sources can contribute considerable fluxes of P (Jarvie et al. 2013; Haygarth et al. 2014; Chen et al. 2015a) and nitrogen (Van Meter and Basu 2017; Kolbe et al. 2019; Pinay et al. 2018) to downstream waterbodies. These legacy nutrient sources can delay water quality recovery for decades or even centuries after cessation of new nutrient inputs (Carpenter 2005; Sharpley et al. 2013; Van Meter and Basu 2017).

Previous studies posit that legacy P stocks are mainly regulated by watershed P buffering capacity, i.e., the ability of a watershed to modulate P flux to downstream waters by retaining excess P within watersheds and modulating P release to receiving waters (Doody et al. 2016; Kusmer et al. 2018; Withers et al. 2018). P buffering capacity is influenced by soil factors (e.g., soil P content, texture, chemistry and mineralogy, Kleinman et al. 2011), hydrological factors (e.g., hydrologic flowpaths, water yield and baseflow index, Kusmer et al. 2018), landscape factors (e.g., topography, land use/land cover, Qiu and Turner 2015), and human and management factors (e.g., artificial soil drainage, river channelization and riparian zone management, Reed and Carpenter 2002; Gentry et al. 2007). Although watershed P buffering capacity is gaining increased attention in developed nations (Abbott et al. 2018; Gu et al. 2019), there is a lack of information concerning P buffering capacity and legacy P dynamics for watersheds in developing nations. Compared to decreasing P inputs commonly observed in watersheds of developed nations, large net P accumulations are more frequently reported in developing nations due to the rapidly increasing use of P fertilizer in recent decades (Haygarth et al. 2014; Jiang and Yuan 2015; Powers et al. 2016; Chen et al. 2018). Therefore, it is highly warranted to examine long-term legacy P stocks and riverine P flux dynamics as mediated by P buffering capacity to achieve P balance in developing countries and avoid/remediate persistent P water pollution issues commonly occurring in developed countries (Sharpley et al. 2013; Chen et al. 2018).

Based on the long-term quantitative relationship between legacy P stocks and riverine P flux data accumulated over 110 years in 23 watersheds of the St Lawrence Basin, Goyette et al. (2018) quantified the legacy P stock threshold (beyond which small increments in P accumulation result in large increases in P delivery to surface waters) and predicted the time lag involved in a return to baseline conditions if anthropogenic P inputs ceased. However, this approach has not been extensively applied due to the lack of long-term records for P inputs and riverine transport in many regions. Importantly, in practice it is not necessary to reduce legacy P stocks and control P levels in surface waters to natural background levels in intensively-managed watersheds. For example, it is also necessary to maintain soil P levels within an optimum range for producing high crop yields (Sharpley et al. 2013; Rowe et al. 2016). Referring to common water pollution control plans, i.e., total maximum daily load (TMDL) (USEPA 2006), the management goal is to determine the critical legacy P stock that provides attainment of water quality standards and to predict how long it will take for the contemporary legacy stock to return to this critical level. Such information allows watershed managers to formulate quantitative targets and schedules for P management plans through regulating source P inputs and legacy P stocks to achieve a target riverine P concentration.

This study examined long-term (1960–2010) accumulated legacy P stocks, P buffering capacity and riverine TP flux dynamics in the Yongan watershed of eastern China. The Yongan watershed is a typical watershed in eastern China, where watersheds generally have small to medium drainage areas (e.g., 100–10,000 km^2) with sub-tropical climate and highly weathered and acidic soils (Oxisols and Ultisols), and have experienced rapid population growth and economic development in recent decades (Chen et al. 2015a, 2016). By integrating the net anthropogenic P input approach (Russell et al. 2008) with long- and short-term buffering indexes (Kusmer et al. 2018), we examined changes in the long-term watershed P buffering capacity and its response to changes in land use/land cover and climate. A new model that incorporates accumulated legacy P stocks and annual precipitation was then developed to estimate the critical legacy P stock threshold necessary to achieve a targeted water quality standard. Finally, an exponential decay model was formulated to predict the time required for depletion of the accumulated legacy P stock via riverine export to meet target riverine P concentration objectives. This is the first attempt to provide a quantitative assessment of long-term P buffering and legacy P stock dynamics at the watershed scale for management practices typical of developing countries.

Materials and methods

Watershed characteristics

The Yongan watershed (120.23–121.01°E and 28.47–29.04°N; elevation ~ 15–1000 m above mean sea level) is located in eastern Zhejiang Province, China (Fig. 1). The Yongan River is the third largest river of Zhejiang Province and flows into Taizhou Estuary and the East China Sea, a coastal area that commonly experiences hypoxia (Gao and Zhang 2010). The river drains a total area of 2474 km^2 and has an average annual water depth of 5.42 m and discharge of 72.9 m^3 s^{-1} at the sampling location. Climate is subtropical monsoon having an average annual temperature of 17.4 °C and average annual precipitation of 1400 mm. Rainfall mainly occurs in May–September (~ 62% of annual precipitation) with a typhoon season in July–September, while winter (December–February) is a major dry season (~ 12%). Although no significant trends in annual precipitation amount or average river discharge occurred over the 1980–2010 period, this watershed experienced a ~ 90% increase in the number of annual storm events (> 50 mm per 24 h, Chen et al. 2015a, 2016).

Agricultural land (including paddy field, garden plot, and dryland) averaged ~ 12% of total watershed area in 1960–2010, with developed land (including rural and urban residential lands, roads, mining and industry), woodlands, and natural lands (including waterbodies, swamp, rock, and natural conservation land) contributing ~ 3%, ~ 67%, and ~ 18%, respectively (Table 1). The watershed is dominated by highly weathered and acidic soils, i.e., Oxisols and Ultisols (Chen et al. 2015a). The economic role of agriculture has been increasingly replaced by industry since the 1990s, resulting in a large reduction in chemical P fertilizer application (~ 38%) and cultivated crop area (~ 20%) since 2000 (Table 1). In 1960, 96% of animal and human solid waste was recycled on agricultural lands; this figure declined to 10% in 2010 (Agricultural Bureau of Xianju County 2011). Agricultural land area irrigated and drained with cement channels and pipes decreased between 1960 and 1980 due to damage, while rose by 91% since 1980 due to renovating and rebuilding. Total population within the watershed increased from ~ 380,000 (97% rural and 3% urban) in 1960

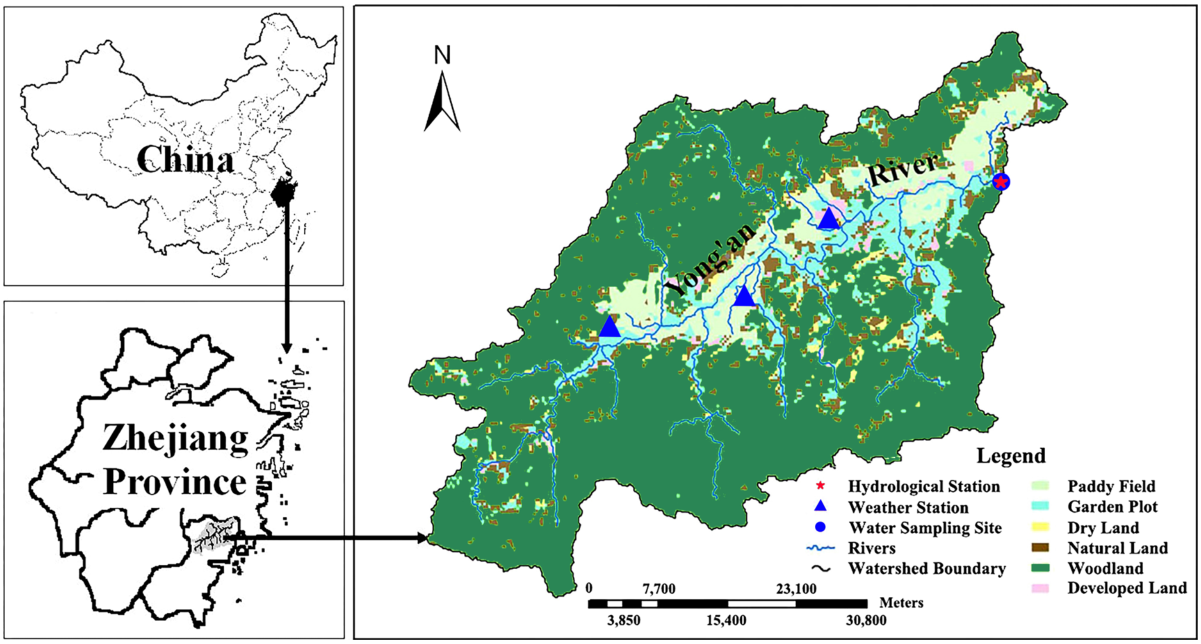

Fig. 1 Land use/land cover in the Yongan watershed with the water quality sampling site and three weather stations designated. (Color figure online)

Table 1 Characteristics of land-use distribution, population, domestic animal, and waste management for the Yongan watershed over the 1960–2010 period

Periods	1960s	1970s	1980s	1990s	2000s
Precipitation (mm year^{-1})	1342	1411	1387	1403	1418
Water discharge (m^3 s^{-1})	67.2	72.4	78.3	72.8	70.0
Agricultural land (%)	12	11	11	11	13
Residential land (%)	2	2	2	3	3
Forest (%)	68	68	68	68	67
Barren land (%)	18	19	18	18	17
Annual area planted to crops (km^2)	598	671	637	601	470
Percentage of agricultural lands with improved drainage (%, e.g., with cement channels and pipes)	30%	25%	19%	18%	29%
Population density (capita km^{-2})	171	222	248	266	289
Livestock density (capita km^{-2})	38	63	100	86	62
Poultry density (capita km^{-2})	13	43	128	226	336
Freshwater aquaculture porduction (ton year^{-1})	1.6	2.2	3.2	13.1	26.8
Recycled animal excreta to croplands (%)	98	97	95	75	57
Recycled rural/urban human excreta to croplands (%)	95/90	93/85	90/80	65/40	25/10
Drained agricultural land area percentage (%)	12	15	19	18	29

The number of each type of livestock and poultry is converted into the number of pigs and chickens according to their phosphorus excretion rates as shown in Chen et al. (2016). All values are the average for each time period

to ~ 740,000 (89% rural and 11% urban) in 2010. Over the 51-year study period, domestic livestock production (pig, cow, sheep, and rabbit) increased ~ 160% during 1960–1980 and decreased ~ 40%

during 1980–2010, while poultry production (chicken and duck) rose 4.8-fold, and freshwater aquaculture production (fish and shrimp) increased 11.7-fold (Table 1).

Net anthropogenic P input (NAPI) estimation and uncertainty analysis

Net anthropogenic P input (NAPI) is a simple yet powerful budget approach to evaluate net P inputs from anthropogenic sources to terrestrial ecosystems (Russell et al. 2008; Han et al. 2013). In general, NAPI (kg P km^{-2} year^{-1}) is calculated as the sum of five major components: atmospheric P deposition (APD), chemical fertilizer P application (CF), non-food input (NFI, e.g., detergent P), seed input (SI), and net food/feed input (NFFI):

$$NAPI = CF + APD + NFI + SI + NFFI \tag{1}$$

Considering the inherent difference in P retention and P delivery efficiencies and pathways between agricultural lands and residential lands, Chen et al. (2016) further divided NAPI into separate P budgets for agricultural ($NAPI_A$) and residential ($NAPI_R$) systems:

$$NAPI = NAPI_A + NAPI_R \tag{2}$$

where NAPI is the sum of $NAPI_A$ and $NAPI_R$. $NAPI_A$ was calculated as the sum of atmospheric deposition, fertilizer application (including chemical and manure fertilizers), and seed P input minus the sum of crop production P and P in residue used for feed (Chen et al. 2016). $NAPI_R$ was calculated as the sum of human and livestock P consumption minus the sum of P contained in livestock production and human/livestock excrement applied to croplands as manure fertilizer (Chen et al. 2016). Detailed descriptions of individual P sources and sinks as well as the parameters used in estimating $NAPI_A$ and $NAPI_R$ are available in Chen et al. (2015a, 2016).

In this study, NAPI, $NAPI_A$ and $NAPI_R$ were estimated for the period 1960–2010. Data for P sources and sinks used to determine P budgets, land-use, domestic waste treatment, and recycled animal/human waste for fertilizing croplands were derived from annual local statistics yearbooks and an investigation of rural ecological and environment quality conducted by the local agricultural bureau (Agricultural Bureau of Xianju County 2011). To gain insight into the uncertainty of $NAPI_A$ and $NAPI_R$ estimates, an uncertainty analysis was performed using Monte Carlo simulation. Due to the limited number of parameter values available in the literature for $NAPI_A$ and $NAPI_R$ estimations, it is difficult to directly determine the probability distribution type and statistical characteristics (i.e., mean and standard deviation) for each parameter. In performing the Monte Carlo simulation for this study, we assumed that all parameters used in $NAPI_A$ and $NAPI_R$ estimates followed a normal distribution with a coefficient of variation of 30% for each parameter, an approached that has been applied in watershed nutrient budgeting studies conducted in nearby regions (Yan et al. 2011; Ti et al. 2012). A total of 10,000 Monte Carlo simulations were performed to obtain the mean and 95% confidence interval for annual $NAPI_A$ and $NAPI_R$.

Annual accumulated legacy P stock was estimated as the sum of accumulated NAPI between 1960 and the year of interest minus the sum of riverine P export flux between 1960 and 1 year before the year of interest:

$$AP_i = \sum_{i=1960}^{n} (NAPI_i - F_{i-1}) \tag{3}$$

where AP_i is accumulated legacy P stock in the year of interest (kg P km^{-2}) and F_{i-1} is riverine TP flux in the year before the year of interest. It should be pointed out that accumulated legacy P stock did not include total P losses to surface waters before 1980. Therefore, the final accumulated legacy P stock in 2010 was estimated as the sum of accumulated NAPI since 1960 minus the sum of total riverine P flux since 1980. Based on riverine P loads in the early 1980s ($\sim$ 10 kg P km^{-2} year^{-1}), we assume that the legacy P stock was overestimated by no more than 200 kg P km^{-2} or < 1% of the calculated legacy P for the study period.

Riverine TP export flux estimates

Daily discharge and river water quality data were not available for the Yongan watershed before 1979, which is common for many rivers in China. Over the 1980–2010 period, water samples (n = 183) were collected once every 4 to 8 weeks at the watershed outlet (Fig. 1). Data for river TP concentration (measured in non-filtered water samples following persulfate digestion using the ammonium molybdate spectrophotometric method, Chen et al. 2015a) and daily discharge were obtained from the local Environmental Protection Bureau and Hydrology Bureau, respectively. To determine whether the daily discharge and TP load time series data were stationary through time, the Augmented Dickey–Fuller Test method was adopted (Hamilton 1994). Results showed that both daily discharge (− 29.65) and TP load (− 12.7) time series data were lower than the 1% test critical values (− 3.44 for sampling number > 500, − 3.51 for sampling number > 100, Sjö 2008), indicating the stationarity of these time series data. To estimate annual TP flux based on the discrete TP concentrations for each sampling site, the LOADEST model (Runkel et al. 2004) was applied for predicting daily TP flux, resulting in a high R^2 value ($R^2 = 0.78$, $p < 0.001$) and Nash–Sutcliffe efficiency coefficient (NSE = 0.78) between measured and estimated values (Chen et al. 2015a, 2016). To verify the reliability of estimated riverine TP load and further evaluate the assumption of stationarity of these time series data, the LOADEST model was independently applied to three separate time segments (1980–1989, 1990–1999 and 2000–2010). Although the predicting accuracies for 1980–1989 ($R^2 = 0.90$, NSE = 0.86), 1990–1999 ($R^2 = 0.89$, NSE = 0.89) and 2000–2010 ($R^2 = 0.88$, NSE = 0.88) improved, the daily TP loads predicted for the three discharge time periods were highly consistent to that of the entire time series ($R^2 = 0.98$, NSE = 0.98) with a relative error of ± 10% for annual riverine TP loads. These results indicate that the established LOADEST model is reasonable for estimating daily TP loads in the Yongan watershed. Using the calibrated LOADEST model, daily riverine TP concentration was predicted using daily river discharge to calculate daily TP loads that were summed to determine the annual riverine TP load. Annual TP flux (kg P km^{-2} $year^{-1}$) was determined by dividing the annual TP load by the watershed area. Annual volume-weighted TP concentration was estimated as the ratio between annual load and total river discharge (Voss et al. 2006).

Watershed P buffering capacity estimates

To examine changes in P buffering capacity for the Yongan watershed over time, this study adopted the short-term buffering index (SBI) and long-term buffering index (LBI) proposed by Kusmer et al. (2018). SBI was calculated as the difference between annual NAPI and annual riverine P flux for the corresponding year:

$$SBI = 1 - \frac{\log(\text{Riverine TP flux })}{\log(\text{annual NAPI})} \quad (4)$$

The LBI was calculated as the difference between accumulated NAPI (between 1960 and the year of interest) and riverine P flux in the year of interest:

$$LBI = 1 - \frac{\log(\text{Riverine TP flux })}{\log(\text{cumulative NAPI })} \quad (5)$$

SBI and LBI represent the capacity of the watershed to retain recently applied P and the historically accumulated surplus P, respectively. These two metrics are very similar to the inverse of the fractional P export (i.e., proportion of NAPI exported by rivers) of annual NAPI and multi-year average NAPI by rivers, respectively, which have been utilized in previous studies (Hong et al. 2012; Goyette et al. 2016).

The significance of all regression analyses in this study was determined by *F*-Test using SPSS statistical software (version 17.0, SPSS, Chicago, IL, USA). The best regression forms (e.g., linear, exponential, power functions) were determined according to the best goodness-of-fit with NSE and R^2 values.

Results and discussion

Watershed P accumulation and riverine P export flux

In the Yongan watershed, annual NAPI rapidly increased from 129.8 kg P km^{-2} $year^{-1}$ in 1960 to 927.1 kg P km^{-2} $year^{-1}$ in 2000 (Fig. 2a), followed by a 16% decline to 781.2 kg P ha^{-1} $year^{-1}$ in 2010. Consistent with previous studies (Russell et al. 2008; Han et al. 2011; Hong et al. 2012), annual NAPI was positively related to population density ($R^2 = 0.90$, $p < 0.01$), agricultural land area percentage ($R^2 = 0.36$, $p < 0.01$), and developed land area

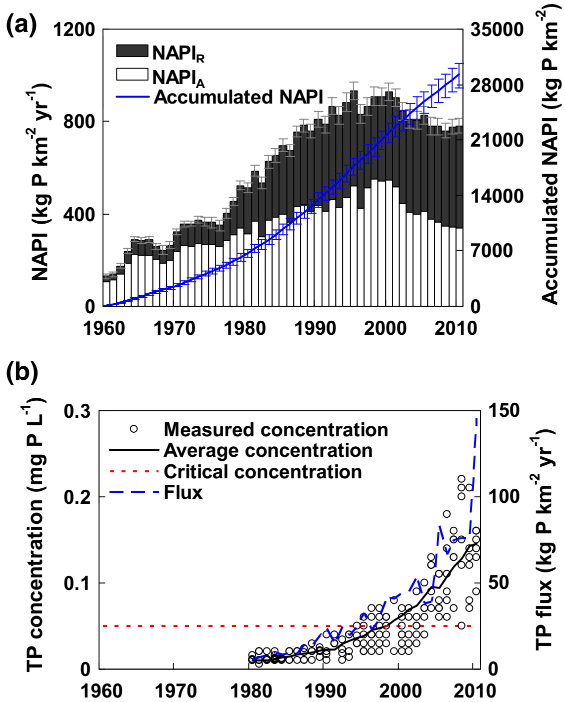

Fig. 2 Historical trends in net anthropogenic phosphorus inputs to entire watershed (NAPI), agricultural land ($NAPI_A$), residential ($NAPI_R$) land, and accumulated legacy P stock (**a**); riverine TP flux and concentration (**b**) in the Yongan watershed over the period 1980–2010. Error bars denote the 95% confidence intervals for NAPI and accumulated legacy P stock from Monte Carlo simulations

percentage (R^2 = 0.89, $p < 0.01$), while it was negatively related to forest area percentage (R^2 = 0.58, $p < 0.01$) over the 1960–2010 period for the entire watershed. Considering the high percentage of forest and natural lands (> 70%, Table 1), estimated NAPI was comparable with the average NAPI estimated for Zhejiang Province, China (i.e., 714–972 kg P km^{-2} year^{-1} in 1981–2009; Han et al. 2013). However, estimated average NAPI (595.7 kg P km^{-2} year^{-1} in 1960–2010) in Yongan watershed was higher than average NAPI reported for mainland of China (~ 465 kg P km^{-2} year^{-1} in 2009, Han et al. 2014), US watersheds (~ 463 kg P km^{-2} year^{-1} in 2007, Han et al. 2012), Baltic Sea watersheds (~ 250 kg P km^{-2} year^{-1} in 1900–2013, McCrackin et al. 2018), and St. Lawrence Basin (~ 230 kg P km^{-2} year^{-1} in 2011, Goyette et al. 2016). In terms of NAPI components for the Yongan watershed, the percentage added to agricultural land ($NAPI_A$) decreased from ~ 80 in 1960 to ~ 43% in 2010, with that added to residential land ($NAPI_R$) increasing from ~ 20 to ~ 57% during the 1960–2010 period. Annual $NAPI_A$ increased 4.0-fold in 1960–2000 and declined 38% over the 2000–2010 period. Annual $NAPI_R$ steadily rose to produce a 43.3-fold increase over the 51-year study period due to a 4.6-fold increase in net food and feed P input and a 150.7-fold increase in non-food P input, as well as an 88% decline in recycled animal/human waste for fertilizing croplands (Table 1).

Over the 1980–2010 period, annual riverine TP flux at the watershed outlet continuously increased by 15-fold on average from 5.71 kg P km^{-2} year^{-1} (0.01 mg P L^{-1}) in 1980 to 145.3 kg P km^{-2} year^{-1} in 2010 (0.15 mg P L^{-1}, Fig. 2b). Since 1998, the river has exceeded the critical concentration of 0.05 mg P L^{-1} that has been established as the maximum annual average concentration to reduce the risk of excessive algal growth (Li et al. 2010). Although the riverine TP flux showed a significant positive relationship with annual NAPI (Fig. 3a), riverine TP flux increased more rapidly than NAPI in 1980–2010 (15.0-fold vs. 1.7-fold increase). Furthermore, the positive relationship between riverine TP flux and NAPI became stronger when the annual NAPI was replaced with the accumulated NAPI (R^2 = 0.42 vs. R^2 = 0.94, Fig. 3). The riverine TP flux also showed a significant positive relationship with precipitation (R^2 = 0.15, $p < 0.01$), number of storm events (> 50 mm per 24 h, R^2 = 0.32, $p < 0.01$), drained agricultural land percentage (R^2 = 0.48, $p < 0.01$), and developed land percentage (R^2 = 0.45, $p < 0.01$). Increases of these influencing factors enhance P transport efficiency from watershed landscapes to rivers via enhanced soil runoff/erosion and sediment transport as well as direct wastewater discharge, decreased water residence time, increased P leaching, and decreased stream sedimentation (Borbor-Cordova et al. 2006; Hong et al. 2012; Morse and Wollheim 2014; Chen et al. 2016).

In 1980–2010, annual NAPI largely exceeded riverine P fluxes (Fig. 2), implying that a considerable fraction of annual NAPI was accumulated in watershed landscapes. This gap in the P balance represents P inputs that are buffered by watershed processes for short to long time periods. Estimated net accumulated NAPI (after riverine P flux was deducted) increased by 306-fold on average from 129.8 kg P km^{-2} (95% CI 118.1–141.7 kg P km^{-2}) in 1960 to 29,258 kg P km^{-2} (95% CI 27,467–30,449 kg P km^{-2}) in 2010 (Fig. 2a), suggesting that > 90% of annual NAPI accumulated

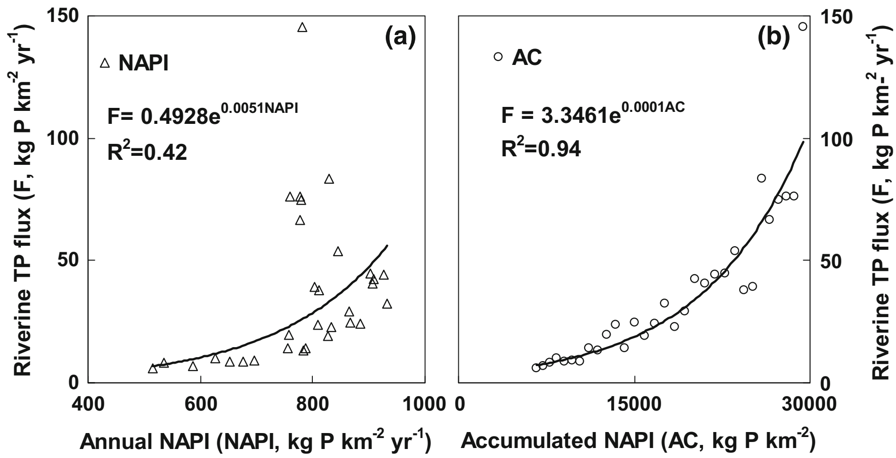

Fig. 3 Relationships between riverine TP flux (F) and annual net anthropogenic phosphorus inputs (NAPI) (**a**) and accumulated legacy P stock (AC) (**b**) in the Yongan watershed over the 1980–2010 study period. The equations and R^2 values represent the regressions and their determination coefficients between riverine TP flux and annual NAPI or accumulated NAPI

as legacy P in the watershed. Estimated average accumulated NAPI in the Yongan watershed ($\sim$ 573.7 kg P km^{-2} year^{-1} in 1960–2010) is higher than that observed in the Thames River watershed ($\sim$ 290 kg P km^{-2} year^{-1} in 1936–2010, Powers et al. 2016), Maumee River watershed ($\sim$ 430 kg P km^{-2} year^{-1} in 1976–2010, Powers et al. 2016), Baltic Sea watershed ($\sim$ 230 kg P km^{-2} year^{-1} in 1900–2013, McCrackin et al. 2018) and St. Lawrence Basin ($\sim$ 220 kg P km^{-2} year^{-1} in 1901–2011, Goyette et al. 2018). A previous analysis indicated that average TP content in the upper 20 cm of agricultural soils increased 2.0-fold between 1984 and 2009 in the Yongan watershed, resulting in a net P accumulation of $\sim$ 8700 kg P km^{-2} year^{-1} (Chen et al. 2016), which accounted for 42% of total NAPI and 91% of total $NAPI_A$ between 1984 and 2009. Phosphorus storage in wetlands, riparian buffers, residential soils, and sediments are likely sinks for the remaining accumulated NAPI (Chen et al. 2015a). In the Yongan watershed, accumulated legacy P stocks explained a greater fraction of the annual riverine TP flux variability than the year-by-year series (Fig. 3). Furthermore, riverine TP fluxes continuously increased in spite of a decline in NAPI between 2000 and 2010 (Fig. 2). These results imply a considerable contribution of legacy P sources as well as an increased fractional export of P inputs to rivers corresponding to decreased buffering capacity over time (Jarvie et al. 2014; Sharpley et al. 2013; Kusmer et al. 2018).

Watershed P buffering capacity dynamics

In the Yongan watershed, the long-term P buffering index (LBI) showed a progressive decrease of 36% from 0.80 in 1980 to 0.52 in 2010 (Fig. 4). In comparison, the short-term P buffering index (SBI) decreased more rapidly showing an overall decline of 65% from 0.72 in 1980 to 0.25 in 2010. The SBI value is an indication of the short-term trapping/retention ability of recently applied P within the watershed, while the LBI value reflects the capacity of the watershed to retain historically accumulated surplus P (Kusmer et al. 2018). The LBI was higher than SBI (Fig. 4), potentially highlighting stronger retention by biophysical storage mechanisms for LBI compared to physical storage mechanisms for SBI (Kusmer et al. 2018). From a long-term perspective, surplus anthropogenic P experiences a range of transport and transformation processes over time and space that

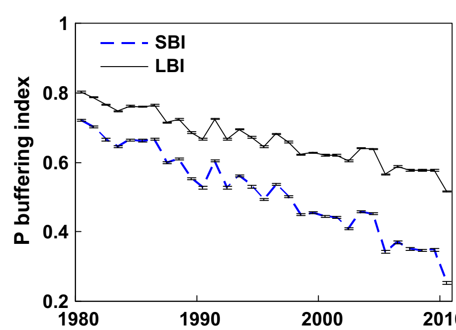

Fig. 4 Historical trends in short-term (SBI) and long-term (LBI) P buffering indexes in the Yongan watershed over the 1980–2010 study period. Error bars denote the 95% confidence intervals for SBI and LBI

result in the potential for longer term retention in biomass (i.e., wetlands and riparian buffers), deeper soil layers, and groundwater reservoirs (Chen et al. 2018). With respect to short-term (within 1 year) buffering processes, transport and transformation of excess P is expected to be mainly regulated by surface terrestrial attributes (i.e., vegetation status) related primarily to runoff/erosion/deposition dynamics (Kusmer et al. 2018). Declining trends for both LBI and SBI were primarily related to the increasing accumulated legacy P stock, and changes in land use and hydroclimate (Table 2, Doody et al. 2016; Kusmer et al. 2018).

Decreased SBI and LBI were associated with increasing accumulated legacy P stock and NAPI (Table 2), implying a progressive P saturation of the watershed with increasing P retained and added (Kusmer et al. 2018; Goyette et al. 2018). In addition, SBI and LBI values both tended to increase as forest and natural land areas increased. Previous studies found forest and natural lands (i.e., water surface, wetlands, and natural conservation lands) to be important landscapes for retaining P via biological, chemical and physical processes (Jiang and Yuan 2015; Cui et al. 2015). In contrast, both SBI and LBI decreased as developed land area increased. Expanding developed land area decreases P retention capacity and enhances flushing (i.e., runoff/erosion) during rainfall events due to a greater impervious surface area that limits infiltration/retention (Chen et al. 2015a).

Table 2 Results of regression analysis between short-term (SBI) or long-term (LBI) P buffering indexes and various independent variables (x) in the Yongan watershed in 1980–2010

Variables	SBI	LBI
Accumulated legacy P stocks (kg P km^{-2} $year^{-1}$)	$SBI = 78.594x^{-0.5224}$	$LBI = 6.3445x^{-0.2329}$
	$R^2 = 0.84^{**}$	$R^2 = 0.89^{**}$
NAPI (kg P km^{-2} $year^{-1}$)	$SBI = 191.83x^{-0.8957}$	$LBI = 14.578x^{-0.4646}$
	$R^2 = 0.28^{**}$	$R^2 = 0.40^{**}$
Forest (%)	SBI = 8E–06 exp(16.518x)	LBI = 0.0095exp(6.3663x)
	$R^2 = 0.53^{**}$	$R^2 = 0.42^{**}$
Natural conservation land (%)	SBI = 0.0213exp(16.757x)	LBI = 0.17exp(7.2421x)
	$R^2 = 0.22^{**}$	$R^2 = 0.22^{**}$
Developed land (%)	SBI = 1.6425exp(− 24.127x)	LBI = 1.0632exp(− 9.5163x)
	$R^2 = 0.59^{**}$	$R^2 = 0.49^{**}$
Agricultural land (%)	$SBI = 0.0008x^{-3.0839}$	$LBI = 0.0521x^{-1.2223}$
	$R^2 = 0.70^{**}$	$R^2 = 0.59^{**}$
Percentage of agricultural lands with improved drainage (%)	SBI = 1.0337exp(− 3.3274x)	LBI = 0.886 exp(− 1.3137x)
	$R^2 = 0.60^{**}$	$R^2 = 0.49^{**}$
Annual area planted to crops (km^2)	$SBI = 4E–05x^{1.4821}$	$LBI = 0.0128x^{0.6232}$
	$R^2 = 0.70^{**}$	$R^2 = 0.65^{**}$
Precipitation (m)	SBI = 1.123exp(− 0.581x)	$LBI = 0.749x^{-0.3661}$
	$R^2 = 0.17^{**}$	$R^2 = 0.18^{**}$
Number of storm events (> 50 mm per 24 h)	$SBI = 0.7663x^{-0.1148}$	$LBI = 0.6735x^{-0.2422}$
	$R^2 = 0.33^{**}$	$R^2 = 0.27^{**}$
Water discharge (m^3 s^{-1})	SBI = 0.6081exp(− 0.0027x)	LBI = 0.7373exp(− 0.0014x)
	$R^2 = 0.04$	$R^2 = 0.06$
Baseflow index	SBI = 1.1768x− 0.4501	LBI = 0.7594x + 0.0465
	$R^2 = 0.20^{**}$	$R^2 = 0.24^{**}$
Surface runoff index	SBI = − 1.1768x + 0.7267	LBI = − 0.7594x + 0.8058
	$R^2 = 0.20^{**}$	$R^2 = 0.24^{**}$

**Significant $p < 0.01$

Additionally, more developed land area implies more waste production (Table 1) and greater P input to rivers via wastewater discharge directly into rivers (Sobota et al. 2011). This is especially significant since centralized wastewater collection is limited in the study area ($< 30\%$ with the remainder directly discharged to surface waterbodies and soils, Table 1). Similarly, increased total agricultural land area and agricultural land area drained with cement channels and pipes both showed a significant positive relationship with SBI and LBI (Table 2). Increased total agricultural land area implies higher P fertilizer additions for crop production, resulting in greater surplus P in agricultural lands. Improved drainage infrastructure enhances hydrologic connectivity and P transport from land to rivers (Gentry et al. 2007; Jarvie et al. 2013). The improved agricultural drainage system (e.g., tile and channel drainage system) facilitates both particulate P transport and dissolved P leaching through lateral and preferential flowpaths in soil, which have been identified as important pathways for transporting dissolved legacy P to rivers (Kleinman et al. 2011; Gentry et al. 2007). However, increasing annual area planted to crops (i.e., multiple crops planted in the same agricultural land in different seasons of a given year), which is beneficial for retaining residual P in soils (e.g., increased soil cover decreases runoff/erosion) and enhances P crop uptake (Gaba et al. 2014), would contribute to a increase in SBI and LBI (Table 2).

Watershed hydrological characteristics have a major impact on both short-term and long-term P retention processes (Metson et al. 2017). In the Yongan watershed, SBI and LBI values both decreased with increasing precipitation and number of storm events (> 50 mm per 24 h) (Table 2). Higher precipitation enhances erosion by increasing surface runoff and higher stream velocities remobilize sediment-bound P from stream channels, which decrease watershed P buffering capacity (Borbor-Cordova et al. 2006; Morse and Wollheim 2014). Similarly, an increasing number of storm events accelerates P loss by increasing both runoff/erosion and leaching, especially from urban runoff and agricultural lands (Shigaki et al. 2007; Sharpley et al. 2008). This is particularly significant for soils that have received high rates of P amendments for long periods of time, since P retention is inversely related to the saturation of a soil's P sorption capacity, especially for the topsoil horizon (Kleinman et al. 2011; Sharpley et al. 2013). Enhanced erosion was documented in the Yongan watershed by a 20% increase in total suspended solid concentration during the high flow regime between 1980 and 2010 (Chen et al. 2015b). The annual baseflow (73–91%) and surface runoff (9–27%) components were estimated in a previous study with the Regional Nutrient Management (ReNuMa) model, a hydrologically-driven, quasi-empirical model is designed to estimate runoff and groundwater contributions as well as nutrient loads in mesoscale watersheds (Hu et al. 2018). Results showed that SBI and LBI were both positively related to the baseflow index and negatively related to the surface runoff index (Table 2). The hydrologic flowpath is a dominant factor as P leaching through the soil-vadose zone-groundwater pathway to produce baseflow is a slow process (months to decades) compared to P lost from soils through surface runoff/erosion (hours to months, Sanford and Pope 2013). Thus, years with a higher baseflow index resulted in greater P buffering capacity. Although precipitation was strongly related to river water discharge ($Q = 39.204P^{1.8001}$, $R^2 = 0.80$, n = 31), SBI and LBI showed no significant relationship with river discharge due to the contrasting impacts of surface runoff versus baseflow on P buffering capacity.

Modeling riverine TP fluxes and sources

The analyses reported above indicate that increasing legacy P stock (Fig. 2) and decreasing P buffering capacity (Fig. 4) due to changes in land-use and hydroclimate are the main causes for the rapid increase in riverine TP flux in 1980–2010. Building on previous modeling efforts (Chen et al. 2015a, 2016), the following model was developed to predict annual riverine TP flux (F_i, kg P km^{-2} year^{-1}) over the 1980–2010 period:

$$F_i = \alpha P_i^{\beta} \exp(\gamma AP_i) \tag{6}$$

where P_i represents annual precipitation (m year^{-1}), AP_i is accumulated legacy P stock (kg P km^{-2}), and α, β, and γ are fitting parameters. Regression analysis was applied to calibrate the unknown parameters after a logarithmic transformation of Eq. (6), resulting in the following model formulation:

$$F_i = 2.269 P_i^{1.146} \exp(1.106 \times 10^{-4} AP_i) \quad (7)$$

The calibrated model Eq. (7) explained 99% of the variation in riverine TP fluxes over the 31-year study period with a NSE of 0.98 between modeled TP flux and LOADEST-estimated values (Fig. 5a). Considering the complexities of P biogeochemistry and transport along the land–water continuum, model performance was efficient and effective in comparison with several more complex physical-based models [e.g., SWAT, AGNPS, and HSPF, reviewed by Moriasi et al. (2007)]. Importantly, model formulation Eq. (7) presented the highest accuracy for predicting riverine TP fluxes compared to previous NAPI based models developed for the Yongan watershed (Chen et al. 2015a and 2016, R^2 = 0.94–0.96), as well as other watersheds (Russell et al. 2008; Hong et al. 2012, $R^2 < 0.80$). This analysis clearly highlights the importance of considering accumulated legacy P stocks in model construction.

Interpretation of Eq. (7) indicates that the component of riverine TP export flux that is not explained by anthropogenic P sources (i.e., when setting AP_i as zero), represents the contribution of natural background sources (e.g., rock weathering). The remaining component originates from anthropogenic P sources (i.e., legacy P stock and current year's NAPI). Setting AP_i as zero, we estimated natural background sources contributed from 2.50 to 5.29 kg P km^{-2} year^{-1} with an average of 3.68 kg P km^{-2} year^{-1} (equivalent to a riverine concentration of 0.004 mg P L^{-1}), which accounted for 4–48% of annual riverine TP fluxes over the study period (Fig. 5b). This estimate of natural background flux falls within the range of baseline TP fluxes (1–10 kg P km^{-2} year^{-1}) reported in previous studies (Withers and Jarvie 2008; Han et al. 2011; Russell et al. 2008; Hong et al. 2012). The dominant soils in the Yongan watershed are strongly weathered Oxisols and Ultisols (~ 90% of watershed area, Chen et al. 2015a). These strongly weathered soils are believed to contribute a very limited background P flux via chemical weathering. In comparison, anthropogenic P sources (i.e., legacy P stock and current year's NAPI) contributed TP fluxes accounted for 52–96% of annual riverine TP flux (Fig. 5b). Our previous studies indicated that legacy P sources contributed 13–32% of annual riverine TP flux during 1980–2010 in the Yongan watershed, while current NAPI might contribute to the remaining 29–64% of riverine TP flux (Chen et al. 2015a, 2016). Although accumulated legacy P (~ 29,300 kg P km^{-2} year^{-1}) greatly exceeds recent NAPI (~ 800 kg P km^{-2} year^{-1}), current year's NAPI has a higher contribution to riverine P than legacy P sources, which is consistent with a more rapid decline of SBI than LBI observed in 1980–2010 (Fig. 4). Similarly, the rapid increase in point source pollution between 1980 and 2010 (~ 27-fold) would disproportionately affect riverine TP flux (Chen et al. 2015b).

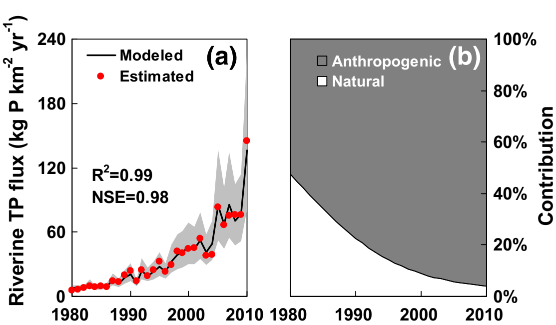

Fig. 5 LOADEST-estimated versus Eq. (7) modeled riverine TP fluxes (**a**) and estimated contributions of natural background and anthropogenic P sources to annual riverine TP flux (**b**) in the Yongan watershed in 1980–2010. Shadow area in (**a**) denotes the 95% confidence interval of the modeled TP fluxes. The determination coefficient (R^2) and Nash–Sutcliffe efficiency coefficient (NSE) values represent the goodness-of-fit between LOADEST-estimated and Eq. (7) modeled riverine TP fluxes

Recovery times required for the depletion of legacy P stock

Following the approach proposed by Goyette et al. (2018), this study estimated the threshold level of accumulated legacy P stock associated with the targeted riverine P concentration of 0.05 mg P L^{-1}. Consistent with the TMDL approach (usually estimated under critical low-flow conditions, e.g., 7Q10, USEPA 2006), the critical riverine P flux was estimated as the product of the critical concentration of 0.05 mg P L^{-1} and the 10th percentile of water discharge (i.e., 49.8 m^3 s^{-1} for Yongan watershed). The critical legacy P stock was inversely estimated from Eq. (7) based on the estimated critical riverine P flux of 31.7 kg P km^{-2} year^{-1} and precipitation of 1.145 m year^{-1} (estimated from the relationship: $Q = 39.204 P^{1.8001}$, R^2 = 0.80, n = 31). The estimated critical legacy P stock was 22.2 ton P km^{-2} (95% CI 19.4–25.3 ton P km^{-2}) for the Yongan watershed, which corresponded to the average accumulated NAPI for the 1998–2001 time period (Fig. 2a). This critical P stock value was tenfold higher than the estimate for the St Lawrence Basin (2.1 ton P km^{-2}, 95% CI 0.03–8.7 ton P km^{-2}, Goyette et al. 2018). This is not surprising because Goyette et al. (2018) estimated the threshold level of legacy P stock based on the breakpoint where P delivery to surface water increased substantially, i.e., where the riverine TP flux exceeded $\sim$ 10 kg P km^{-2} year^{-1}. In terms of the relationship established between accumulated NAPI and riverine TP flux for the St Lawrence basin (annual precipitation ranging from 985 to 1347 mm year^{-1}), the threshold for the accumulated legacy P stock should range between 10 and 25 ton P km^{-2} at a critical riverine TP flux of 31.7 kg P km^{-2} year^{-1}, which is close to our estimate for the Yongan watershed (annual precipitation ranging from 1070 to 1813 mm year^{-1}). These comparisons indicated that estimated critical riverine TP fluxes vary as a function of river discharge, resulting in variable estimates for thresholds of accumulated legacy P stock. Furthermore, the high Al/Fe oxide/hydroxide content of the highly weathering Oxisols and Ultisols in the Yongan watershed may be expected to have a considerably higher P buffering capacity than soils in the St. Lawrence basin (Uehara and Gillman 1981).

We further modeled the time required for depletion of the watershed legacy P stock to reach the estimated critical legacy P stock as an exponential decay process (Goyette et al. 2018): $S(t) = S_0\exp(-Et)$, where S_0 is the P stock in 2010, E is the P export coefficient (year^{-1}) and t is the number of years over which legacy P stock is depleted via riverine export while assuming no additional P inputs. In terms of this exponential decay equation, required recovery time was mainly influenced by NAPI and P accumulation levels, as well as hydroclimate. To express a general scenario, the P export coefficient E value (varying between 8.487 $\times$ 10^{-4} and 4.955 $\times$ 10^{-3} year^{-1}) was set as the average value (1.759 $\times$ 10^{-3} year^{-1}) from the relationship between riverine TP flux and legacy P stock in 1980–2010. The resulting estimate for the time required to deplete accumulated legacy P pools back to the critical level was 159 years (95% CI 57–262 years, Fig. 6a). This indicates that the legacy P stock built up over a relatively short time frame ($\sim$ 13 years between 1998 and 2010) would take several decades to centuries to return to the critical level even with no additional P inputs. To demonstrate the influences of accumulated P legacy stock and climate change on the recovery times, we further predicted the recovery times based on the scenarios of decreasing NAPI and changing climate. For the scenario of a 4% reduction of annual NAPI since 2010 according to the average decreasing rate in 2000–2010, estimated time required for accumulated legacy P pools to return to the critical level was 456 years (95% CI 353–560 years, Fig. 6b). For the scenario of coupling a 4% reduction of annual NAPI and 10% increase in average annual precipitation (compared to average annual precipitation amount in 1980–2010, Kang 2008), estimated recovery time was 318 years (95% CI 238–400 years, Fig. 6c). These predicted results indicated that the recovery times would be highly impacted by climate change components and accumulated P legacy stock.

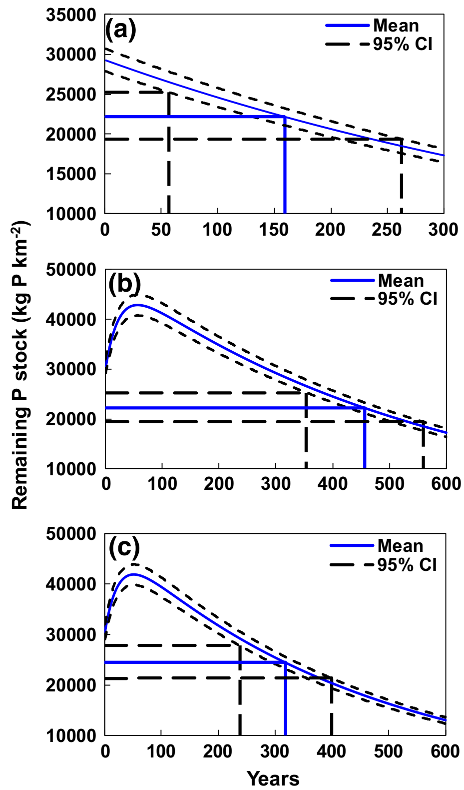

Fig. 6 Recovery time required to deplete the legacy P stock since 2010 to the riverine critical TP level of 0.05 mg P L^{-1} using an exponential decay model under scenarios of **a** no additional P inputs; **b** 4% reduction in annual NAPI and **c** 4% reduction in annual NAPI and 10% increase in average annual precipitation

Although surprisingly long, this time frame is supported by field studies that demonstrated that a decade or more of "P draw down" from agricultural soil P reserves was required to substantially reduce P in runoff even with reduced NAPI or no additional fertilizer application (Sharpley et al. 2013; Chen et al. 2018). Several watershed studies have similarly shown a lag time of several decades to recover from P pollution as reviewed by Meals et al. (2010), Sharpley et al. (2013) and Chen et al. (2018). A key component of these legacy effects is the long hydrological residence times, which can range from years to decades in many watersheds (Meals et al. 2010; Hamilton 2012; Sebilo et al. 2013; Kolbe et al. 2016; Marçais et al. 2018). Consequently, nutrients carried along various hydrological pathways would be potentially delayed by years and decades, resulting in accumulation of legacy nutrient pools in soils, sediments and groundwater. While hydrological residence times for various hydrologic flowpath components are not available in the Yongan watershed, the high baseflow index (0.82 on average) estimated in our previous modeling study implies a potentially long hydrological residence time associated a dominance of groundwater flowpaths (Hu et al. 2018). Watersheds with a higher baseflow component usually have longer hydrological residence times resulting in large legacy pools in the groundwater reservoir and long lag times for dissolved nutrient (e.g., NO_3^- and PO_4^{3-}, Kolbe et al. 2019). However, particulate-associated nutrients can also take years to be transported downstream as particles are repeatedly deposited/trapped, resuspended, and redeposited within the terrestrial-aquatic continuum by episodic high flow events. This process can delay particulate nutrient transport from upstream drainage areas to the downstream outlet by years or even decades, which is dependent on watershed attributes (e.g., area, shape, topography, and land-use/land cover, Pinay et al. 2018; Van Meter and Basu 2017). In terms of previous studies, legacy stock size and recovery time length for dissolved nutrients are highly dependent on the baseflow component and soil properties (e.g., organic matter and nutrient status), while particulate nutrient transport is strongly associated with frequency of high flow events and watershed attributes (Gu et al. 2019; Chen et al. 2018).

It should be pointed out that estimated recovery times required to return to a critical riverine P level imply considerable uncertainties because the model ignores stable storage of P (e.g., strongly occluded P), future climate change (e.g., frequency and intensity of precipitation) and land use/land cover (Sattari et al. 2012; Chen et al. 2015a, 2016; Goyette et al. 2018). A particular uncertainty associated with this modeling approach is accounting for the potentially strong hysteresis effects that are prominent in P sorption/desorption dynamics (i.e., irreversible P desorption characteristics, Okajima et al. 1983; Dubus and Becquer 2001). Clearly, examining watershed-scale P retention/release hysteresis dynamics remains a critical future research topic required to refine estimates of recovery times associated with depletion of legacy P stocks.

Implications for watershed P management

Quantitative information on long-term P buffering capacity dynamics is important for optimizing the delivery of diverse ecosystem services associated with P in watershed management (Doody et al. 2016; Macintosh et al. 2019). In the Yongan watershed, both short- and long-term P buffering indexes continually decreased (Fig. 4) due to an increase of accumulated legacy P stocks (Fig. 2a) and changes in land use and climate (Table 2). As a result, there was a corresponding rapid increase in riverine TP flux that exceeded the critical riverine TP concentration of 0.05 mg P L^{-1} (Fig. 2b). Since it is not feasible to regulate hydroclimate conditions and costly to decrease agricultural and developed land areas, P source input controls, transport interception measures and accumulated P legacy stock reductions are particularly necessary to improve watershed P buffering capacity for a timely and efficient reduction of P loading to surface waters.

Given the more rapid decline of the short-term P buffering index than the long-term P buffering index (Fig. 4), current year's NAPI had a higher contribution (29–64%) to riverine TP fluxes than legacy P sources (13–32%). Therefore, improving short-term P buffering capacity should be a priority through controlling P source inputs and enhancing interception of P during hydrologic transport. Due to low wastewater collection and treatment in the Yongan watershed (< 15% for rural area and < 60% in urban area, Chen et al. 2016), increasing collection and processing of animal and domestic wastewater could greatly reduce the point source P pollution loading directly to surface waters. Considering the 88% decline in recycled animal/human excreta for fertilizing croplands over the past 51 years (Table 1), which is common in many regions of China (Hou et al. 2013; Jiang and Yuan 2015), it would be advantageous to promote recycling of human and domestic animal solid wastes to replace chemical fertilizer use on croplands. Although transport control measures (e.g., buffer strips and wetlands) can increase nutrient retention and reduce P loading to surface waters in the short-term (Table 2), they eventually become P saturated with respect to legacy P and eventually lose their P retention capacity or become a source of P to surface waters (Jarvie et al. 2014; Haygarth et al. 2014; Chen et al. 2015a). Therefore, controlling source P inputs by adopting relevant balancing and cycling measures must be a primary strategy for controlling P loading to surface waters in a long-term perspective.

For the purposes of improving long-term P buffering capacity and reducing P source inputs, a key management strategy is to use less P by balancing P fertilizer application with crop requirements while considering the potentially high availability of P contained in soil legacy P pools (Withers et al. 2014; Rowe et al. 2016). Several studies demonstrated that crops could recover legacy P from long-term over-application of fertilizer and manure P with no reduction in crop yields for several years to decades after cessation of P fertilization (Sharpley et al. 2013; Jarvie et al. 2013; Liu et al. 2016; Rowe et al. 2016; Chen et al. 2018). When considering the depletion of legacy P stock via crop uptake with an average coefficient for 2000–2010 (i.e., 4.63×10^{-3} year^{-1}), the return of the contemporary legacy P stock to the critical level would potentially support crop production for $\sim$ 44 years without any external P inputs (Chen et al. 2017). To sustainably recover legacy soil P pools, several practical measures, i.e., breeding more P efficient plants, changing farm management (e.g., increasing the multiple crop index, Gaba et al. 2014), and applying commercial bioinoculant products (Withers et al. 2014; Rowe et al. 2016), are applicable to enhance soil P use by crop production.

Although estimated recovery times for depletion of legacy P stock to meet a critical riverine P target concentration (Fig. 6) incorporate considerable uncertainty, they provide a conservative estimate for developing quantitative management goals and schedules for regulating source P inputs and legacy P stocks at the watershed scale. A relatively short timescale experienced for accumulating such a large legacy P stock ($\sim$ 13 years; Fig. 2a) and a much longer timescale ($\sim$ 150–460 years) required for recovery to the critical legacy P stock level (Fig. 6) highlight that P balancing and cycling strategies should be adopted in advance to avoid accumulating excessive legacy P stocks in watersheds of developing nations pursuing food security through excessive P fertilizer inputs (Jarvie et al. 2014). A recent review documented that large P accumulations over relatively short timeframes are widely observed in various watersheds across China, Europe and USA over the past several decades (Chen et al. 2018), resulting in considerable lag times in the water quality response to contemporary management measures (Meals et al. 2010). These results highlight that a long-term perspective is required for developing sustainable P management strategies to maximize both agronomic and environmental goals (Haygarth et al. 2014; Kusmer et al. 2018).

Conclusion

This study provides new insights into the long-term decline in watershed P buffering capacities in response to growing accumulated legacy P stocks coupled with changing land use and climate, resulting in a rapid increase of riverine P export flux. Aimed at improving watershed P buffering capacity, P source input controls, transport interceptions, and accumulated legacy stock reductions are required for a timely and efficient reduction of P loading to surface waters. Although climate change may shorten recovery times due to increased precipitation/river discharge in the Yongan watershed, excessive legacy P stocks built up over a relatively short time frame are expected to require long time periods to drop back to targeted levels, even with reductions or cessation of new P inputs. We highlight that for P-limited agricultural watersheds, P balancing and cycling strategies should be adopted in advance to avoid accumulation of excessive legacy P stocks that would result in persistent water quality degradation. For P enriched watersheds, controlling P source inputs and utilizing legacy soil P pools should be primary strategies for protecting/remediating water quality and conserving rock P resources. A long-term perspective is required for developing sustainable P management strategies to maximize the agronomic and environmental benefits.

Acknowledgements We thank local government departments for providing data critical for this investigation. This work was supported by the National Natural Science Foundation of China (51679210 and 41877465), National Key Research and Development Program of China (2017YFD0800101) and Zhejiang Provincial Natural Science Foundation of China (LR19D010002).

Compliance with ethical standards

Conflict of interest This study has no conflict of interest with any persons or affiliations.

References

Abbott BW, Moatar F, Gauthier O, Fovet O, Antoine V, Ragueneau O (2018) Trends and seasonality of river nutrients in agricultural catchments: 18 years of weekly citizen science in France. Sci Total Environ 624:845–858. https://doi.org/10.1016/j.scitotenv.2017.12.176

Agricultural Bureau of Xianju County (2011) Agricultural Ecological Environment Quality Assessment in the Yongan River Watershed. http://www.xjlsncp.com/html/main/nyqhView/2011-12/6533.html. Accessed Oct 2014

Borbor-Cordova MJ, Boyer EW, McDowell WH, Hall CA (2006) Nitrogen and phosphorus budgets for a tropical watershed impacted by agricultural land use: Guayas, Ecuador. Biogeochemistry 79:135–161. https://doi.org/10.1007/s10533-006-9009-7

Carpenter SR (2005) Eutrophication of aquatic ecosystems: bistability and soil phosphorus. Proc Natl Acad Sci 102:10002–10005. https://doi.org/10.1073/pnas.0503959102

Carpenter SR (2008) Phosphorus control is critical to mitigating eutrophication. Proc Natl Acad Sci 105:11039–11040. https://doi.org/10.1073/pnas.0806112105

Chen DJ, Hu MP, Guo Y, Dahlgren RA (2015a) Influence of legacy phosphorus, land use, and climate change on anthropogenic phosphorus inputs and riverine export dynamics. Biogeochemistry 123:99–116. https://doi.org/10.1007/s10533-014-0055-2

Chen DJ, Hu MP, Guo Y, Dahlgren RA (2015b) Reconstructing historical changes in phosphorus inputs to rivers from point and nonpoint sources in a rapidly developing watershed in eastern China, 1980–2010. Sci Total Environ 533:196–204. https://doi.org/10.1016/j.scitotenv.2015.06.079

Chen DJ, Hu MP, Wang JH, Dahlgren RA (2016) Factors controlling phosphorus export from agricultural/forest and residential systems to rivers in eastern China, 1980–2011. J Hydrol 533:53–61. https://doi.org/10.1016/j.jhydrol.2015.11.043

Chen DJ, Hu MP, Guo Y (2017) Long-term (1980–2010) changes in cropland phosphorus budgets, use efficiency and legacy pools across townships in the Yongan watershed, eastern China. Agric Ecosyst Environ 236:166–176. https://doi.org/10.1016/j.agee.2016.12.003

Chen DJ, Shen H, Hu MP, Wang JH, Zhang YF, Dahlgren RA (2018) Legacy nutrient dynamics at the watershed scale: principles, modeling, and implications. Adv Agron 149:237–313. https://doi.org/10.1016/bs.agron.2018.01.003

Conley DJ, Paerl HW, Howarth RW, Boesch DF, Seitzinger SP, Havens KE, Lancelot C, Likens GE (2009) Controlling eutrophication: nitrogen and phosphorus. Science 323:1014–1015. https://doi.org/10.1126/science.1167755

Cui SH, Xu S, Huang W, Bai XM, Huang YF, Li GL (2015) Changing urban phosphorus metabolism: evidence from Longyan City, China. Sci Total Environ 536:924–932. https://doi.org/10.1016/j.scitotenv.2015.06.073

Doody DG, Withers PJA, Dils RM, McDowell RW, Smith V, McElarney YR, Dunbar M, Daly D (2016) Optimising land use for the delivery of catchment ecosystem services. Front Ecol Environ 14:325–332. https://doi.org/10.1002/fee.1296

Dubus IG, Becquer T (2001) Phosphorus sorption and desorption in oxide-rich Ferralsols of New Caledonia. Soil Res 39:403–414. https://doi.org/10.1071/SR00003

Gaba S, Lescourret F, Boudsocq S, Enjalbert J, Hinsinger P, Journet E-P, Navas M-L, Wery J, Louarn G, Malézieux E, Pelzer E, Prudent M, Ozier- Lafontaine H (2014) Multiple cropping systems as drivers for providing multiple ecosystem services: from concepts to design. Agron Sustain Dev 35:607–623. https://doi.org/10.1007/s13593-014-0272-z

Gao C, Zhang TL (2010) Eutrophication in a Chinese context: understanding various physical and socio-economic aspects. Ambio 39:385–393. https://doi.org/10.1007/s13280-010-0040-5

Gentry LE, David MB, Royer TV, Mitchell CA, Starks KM (2007) Phosphorus transport pathways to streams in tile-drained agricultural watersheds. J Environ Qual 36:408–415. https://doi.org/10.2134/jeq2006.0098

Goyette JO, Bennett EM, Howarth RW, Maranger R (2016) Changes in anthropogenic nitrogen and phosphorus inputs to the St. Lawrence sub-basin over 110 years and impacts on riverine export. Glob Biogeochem Cycl 30:1000–1014. https://doi.org/10.1002/2016GB005384

Goyette JO, Bennett EM, Maranger R (2018) Low buffering capacity and slow recovery of anthropogenic phosphorus pollution in watersheds. Nat Geosci 11:921–925. https://doi.org/10.1038/s41561-018-0238-x

Gu S, Gruau G, Dupas R, Petitjean P, Li Q, Pinay G (2019) Respective roles of Fe-oxyhydroxide dissolution, pH changes and sediment inputs in dissolved phosphorus release from wetland soils under anoxic conditions. Geoderma 338:365–374. https://doi.org/10.1016/j.geoderma.2018.12.034

Hamilton JD (1994) Time series analysis. Princeton University Press, Princeton

Hamilton SK (2012) Biogeochemical time lags may delay responses of streams to ecological restoration. Freshwater Biol 57:43–57. https://doi.org/10.1111/j.1365-2427.2011.02685.x

Han H, Bosch N, Allan JD (2011) Spatial and temporal variation in phosphorus budgets for 24 watersheds in the Lake Erie and Lake Michigan basins. Biogeochemistry 102:45–58. https://doi.org/10.1007/s10533-010-9420-y

Han H, Allan JD, Bosch NS (2012) Historical pattern of phosphorus loading to Lake Erie watersheds. J Great Lakes Res 38:289–298. https://doi.org/10.1016/j.jglr.2012.03.004

Han YG, Yu XX, Wang XX, Wang YQ, Tian JX, Xu L, Wang CZ (2013) Net anthropogenic phosphorus inputs (NAPI) index application in Mainland China. Chemosphere 90:329–337. https://doi.org/10.1016/j.chemosphere.2012.07.023

Han YG, Fan YT, Yang PL, Wang XX, Wang YJ, Tian JX, Xu L, Wang CZ (2014) Net anthropogenic nitrogen inputs (NANI) index application in Mainland China. Geoderma 213:87–94. https://doi.org/10.1016/j.geoderma.2013.07.019

Haygarth PM, Jarvie HP, Powers SM, Sharpley AN, Elser JJ, Shen JB, Peterson HM, Chan NI, Howden NJK, Burt T, Worrall F, Zhang FS, Liu XJ (2014) Sustainable phosphorus management and the need for a long-term perspective: the legacy hypothesis. Environ Sci Technol. https://doi.org/10.1021/es502852

Hong B, Swaney DP, Mörth C-M, Smedberg E, Hägg EH, Humborg C, Howarth RW, Bouraoui F (2012) Evaluating regional variation of net anthropogenic nitrogen and phosphorus inputs (NANI/NAPI), major drivers, nutrient retention pattern and management implications in the multinational areas of Baltic Sea basin. Ecol Model 227:117–135. https://doi.org/10.1016/j.ecolmodel.2011.12.002

Hou Y, Ma L, Gao ZL, Wang FH, Sims JT, Ma WQ, Zhang FS (2013) The driving forces for nitrogen and phosphorus flows in the food chain of China, 1980 to 2010. J Environ Qual 42:962–971. https://doi.org/10.2134/jeq2012.0489

Howarth R, Chan F, Conley DJ, Garnier J, Doney SC, Marino R, Billen G (2011) Coupled biogeochemical cycles: eutrophication and hypoxia in temperate estuaries and coastal marine ecosystems. Front Ecol Environ 9:18–26. https://doi.org/10.1890/100008

Hu MP, Liu YM, Wang JH, Dahlgren RA, Chen DJ (2018) A modification of the regional nutrient management model (ReNuMa) to identify long-term changes in riverine nitrogen sources. J Hydrol 561:31–42. https://doi.org/10.1016/j.jhydrol.2018.03.068

Jarvie HP, Sharpley AN, Withers PJA, Scott JT, Haggard BE, Neal C (2013) Phosphorus mitigation to control river eutrophication: murky waters, inconvenient truths and 'post-normal' science. J Environ Qual 42:295–304. https://doi.org/10.2134/jeq2012.0085

Jarvie HP, Sharpley AN, Brahana V, Simmons T, Price A, Neal C, Lawlor AJ, Sleep D, Thacker S, Haggard BE (2014) Phosphorus retention and remobilization along hydrological pathways in Karst Terrain. Environ Sci Technol 48:4860–4868. https://doi.org/10.1021/es405585b

Jiang SY, Yuan ZW (2015) Phosphorus flow patterns in the Chaohu watershed from 1978 to 2012. Environ Sci Technol 49:13973–13982. https://doi.org/10.1021/acs.est.5b03202

Kang LL (2008) Prediction of future climate change scenarios in Zhejiang Province. Annual meeting of the Chinese meteorological society: climate change parallel sessions. Yantai City, Shangdong Province, China (**in Chinese**)

Kleinman PJA, Sharpley AN, McDowell RW, Flaten DN, Buda AR, Tao L, Bergstrom L, Zhu Q (2011) Managing agricultural phosphorus for water quality protection: principles for progress. Plant Soil 349:169–182. https://doi.org/10.1007/s11104-011-0832-9

Kolbe T, Marçais J, Thomas Z, Abbott BW, de Dreuzy JR, Rousseau-Gueutin P, Aquilina L, Labasque T, Pinay G (2016) Coupling 3D groundwater modeling with CFC-based age dating to classify local groundwater circulation in an unconfined crystalline aquifer. J Hydrol 543:31–46. https://doi.org/10.1016/j.jhydrol.2016.05.020

Kolbe T, de Dreuzy JR, Abbott BW, Aquilina L, Babey T, Green CT, Fleckenstein JH, Labasque T, Laverman AM, Marçais J, Peiffer S, Thomas Z, Pinay G (2019) Stratification of reactivity determines nitrate removal in groundwater. Proc Natl Acad Sci 116:2494–2499. https://doi.org/10.1073/pnas.1816892116

Kusmer AS, Goyette JO, MacDonald GK, Bennett EM, Maranger R, Withers PJA (2018) Watershed buffering of legacy phosphorus pressure at a regional scale: a comparison across space and time. Ecosystems 22:91–109. https://doi.org/10.1007/s10021-018-0255-z

Li ZW, Yuan J, Bi H, Wu J (2010) Anthropogenic phosphorus flow analysis of Hefei City, China. Sci Total Environ 408:5715–5722. https://doi.org/10.1016/j.scitotenv.2010.08.052

Liu XJ, Vitousek P, Chang YH, Zhang WF, Matson P, Zhang FS (2016) Evidence for a historic change occurring in China. Environ Sci Technol 50:505–506. https://doi.org/10.1021/acs.est.5b05972

MacDonald GK, Bennett EM, Potter PA, Ramankutty N (2011) Agronomic phosphorus imbalances across the world's croplands. Proc Natl Acad Sci 108:3086–3091. https://doi.org/10.1073/pnas.1010808108

Macintosh KA, Doody DG, Withers PJ, McDowell RW, Smith DR, Johnson LT, Bruulsema TW, O'Flaherty V, McGrath JW (2019) Transforming soil phosphorus fertility management strategies to support the delivery of multiple ecosystem services from agricultural systems. Sci Total Environ 649:90–98. https://doi.org/10.1016/j.scitotenv.2018.08.272

Marçais J, Gauvain A, Labasque T, Abbott BW, Pinay G, Aquilina L, Chabaux F, Viville D, de Dreuzy JR (2018) Dating groundwater with dissolved silica and CFC concentrations in crystalline aquifers. Sci Total Environ 636:260–272. https://doi.org/10.1016/j.scitotenv.2018.04.196

McCrackin ML, Muller-Karulis B, Gustafsson BG, Howarth RW, Humborg C, Svanbäck A, Swaney D (2018) A century of legacy phosphorus dynamics in a large drainage basin. Glob Biogeochem Cycle 32:1107–1122. https://doi.org/10.1029/2018GB005914

Meals DW, Dressing SA, Davenport TE (2010) Lag time in water quality response to best management practices: a review. J Environ Qual 39:85–96. https://doi.org/10.2134/jeq2009.0108

Metson GS, Lin JJ, Harrison JA, Compton JE (2017) Linking terrestrial phosphorus inputs to riverine export across the United States. Water Res 124:177–191. https://doi.org/10.1016/j.watres.2017.07.037

Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed

simulations. Trans ASABE 50:885–900. https://doi.org/10.13031/2013.23153
Morse NB, Wollheim WM (2014) Climate variability masks the impacts of land use change on nutrient export in a suburbanizing watershed. Biogeochemistry 121:45–59. https://doi.org/10.1007/s10533-014-9998-6
Okajima H, Kubota H, Sakuma T (1983) Hysteresis in the phosphorus sorption and desorption processes of soils. Soil Sci Plant Nutr 29:271–283. https://doi.org/10.1080/00380768.1983.10434628
Pinay G, Bernal S, Abbott BW, Lupon A, Marti E, Sabater F, Krause S (2018) Riparian corridors: a new conceptual framework for assessing nitrogen buffering across biomes. Front Environ Sci. https://doi.org/10.3389/fenvs.2018.00047
Powers SM, Bruulsema TW, Burt TP, Chan NI, Elser JJ, Haygarth PM, Howden NJK, Jarvie HP, Lyu Y, Peterson HM, Sharpley AN (2016) Long-term accumulation and transport of anthropogenic phosphorus in three river basins. Nat Geosci 9:353–356. https://doi.org/10.1038/ngeo2693
Qiu JX, Turner MG (2015) Importance of landscape heterogeneity in sustaining hydrologic ecosystem services in an agricultural watershed. Ecosphere 6:1–19. https://doi.org/10.1890/ES15-00312.1
Reed T, Carpenter SR (2002) Comparisons of P-yield, riparian buffer strips, and land cover in six agricultural watersheds. Ecosystems 5:568–577. https://doi.org/10.1007/s10021-002-0159-8
Rowe H, Withers PJA, Baas P, Chan NI, Doody D, Holiman J, Jacobs B, Li HG, MacDonald GK, McDowell R, Sharpley AN, Shen JB, Taheri W, Wallenstein M, Weintraub MN (2016) Integrating legacy soil phosphorus into sustainable nutrient management strategies for future food, bioenergy and water security. Nutr Cycles Agroecosyst 104:393–412. https://doi.org/10.1007/s10705-015-9726-1
Runkel RL, Crawford CG, Cohn TA (2004) Load estimator (Loadest): A Fortran program for estimating constituent loads in streams and rivers. http://pubs.usgs.gov/tm/2005/tm4A5/pdf/508final.pdf. Accessed 26 June 2014
Russell MJ, Weller DE, Jordan TE, Sigwart KJ, Sullivan KJ (2008) Net anthropogenic phosphorus inputs: spatial and temporal variability in the Chesapeake Bay region. Biogeochemistry 88:285–304. https://doi.org/10.1007/s10533-008-9212-9
Sanford WE, Pope JP (2013) Quantifying groundwater's role in delaying improvements to Chesapeake Bay water quality. Environ Sci Technol 47:13330–13338. https://doi.org/10.1021/es401334k
Sattari SZ, Bouwman AF, Giller KE, van Ittersum MK (2012) Residual soil phosphorus as the missing piece in the global phosphorus crisis puzzle. Proc Natl Acad Sci 109:6348–6353. https://doi.org/10.1073/pnas.1113675109
Sebilo M, Mayer B, Nicolardot B, Pinay G, Mariotti A (2013) Long-term fate of nitrate fertilizer in agricultural soils. Proc Natl Acad Sci 110:18185–18189. https://doi.org/10.1073/pnas.1305372110
Sharpley AN, Kleinman PJA, Heathwaite AL, Gburek WJ, Folmar GJ, Schmidt JP (2008) Phosphorus loss from an agricultural watershed as a function of storm size. J Environ Qual 37:362–368. https://doi.org/10.2134/jeq2007.0366
Sharpley AN, Jarvie HP, Buda A, May L, Spears B, Kleinman P (2013) Phosphorus legacy: overcoming the effects of past management practices to mitigate future water quality impairment. J Environ Qual 42:1308–1326. https://doi.org/10.2134/jeq2013.03.0098
Shigaki F, Sharpley A, Prochnow LI (2007) Rainfall intensity and phosphorus source effects on phosphorus transport in surface runoff from soil trays. Sci Total Environ 373:334–343. https://doi.org/10.1016/j.scitotenv.2006.10.048
Sjö B (2008) Testing for unit roots and cointegration. https://pdfs.semanticscholar.org/7ce6/2a0c7f6dab85f264a5403bf9b99a0f20a156.pdf. Accessed July 2019
Sobota DJ, Harrison JA, Dahlgren RA (2011) Linking dissolved and particulate phosphorus export in rivers draining California's Central Valley with anthropogenic sources at the regional scale. J Environ Qual 40:1290–1302. https://doi.org/10.2134/jeq2011.0010
Ti CP, Pan JJ, Xia YQ, Yan XY (2012) A nitrogen budget of mainland China with spatial and temporal variation. Biogeochemistry 108:381–394. https://doi.org/10.2307/41410602
Uehara G, Gillman G (1981) The mineralogy, chemistry, and physics of tropical soils with variable charge clays. Westview Press, Boulder
USEPA (2006) An Approach for Using Load Duration Curves in Developing TMDLs. https://www.epa.gov/sites/production/files/2015-07/documents/2007_08_23_tmdl_duration_curve_guide_aug2007.pdf. US Environmental Protection Agency, Washington, DC. Accessed Feb 2019
Van Meter KJ, Basu NB (2017) Time lags in watershed-scale nutrient transport: an exploration of dominant controls. Environ Res Lett 12:084017. https://doi.org/10.1088/1748-9326/aa7bf4
Vitousek PM, Porder S, Houlton BZ, Chadwick OA (2010) Terrestrial phosphorus limitation: mechanisms, implications, and nitrogen–phosphorus interactions. Ecol Appl 20:5–15. https://doi.org/10.1890/08-0127.1
Voss M, Deutsch B, Elmgren R, Humborg C, Kuuppo P, Pastuszak M, Rolff C, Schulte U (2006) Source identification of nitrate by means of isotopic tracers in the Baltic Sea catchments. Biogeosciences 3:663–676. https://doi.org/10.5194/bg-3-663-2006
Withers PJA, Jarvie HP (2008) Delivery and cycling of phosphorus in rivers: a review. Sci Total Environ 400:379–395. https://doi.org/10.1016/j.scitotenv.2008.08.002
Withers PJA, Sylvester-Bradley R, Jones DL, Healey JR, Talboys PJ (2014) Feed the crop not the soil: rethinking phosphorus management in the food chain. Environ Sci Technol 48:6523–6530. https://doi.org/10.1021/es501670j
Withers PJA, Rodrigues M, Soltangheisi A, Carvalho TS, Guilherme LRG, Benites VM, Gatiboni LC, Sousa DMG, Nunes RDS, Rosolem CA (2018) Transitions to sustainable management of phosphorus in Brazilian agriculture. Sci Rep 8:2537. https://doi.org/10.1038/s41598-018-20887-z

Yan XY, Cai ZC, Yang R, Ti CP, Xia YQ, Li FY, Wang JQ, Ma AJ (2011) Nitrogen budget and riverine nitrogen output in a rice paddy dominated agricultural watershed in eastern China. Biogeochemistry 106:489–501. https://doi.org/10.1007/s10533-010-9528-0