Environmental Monitoring and Assessment (2020) 192:54
https://doi.org/10.1007/s10661-019-7967-1

Using maintenance records from a long-term sensor monitoring network to evaluate the relationship between maintenance schedule and data quality

Donald R. Schoolmaster Jr. · Sarai C. Piazza

Received: 8 July 2019 / Accepted: 11 November 2019

Abstract Sensor-based environmental monitoring networks are beginning to provide the large-scale, long-term data required to address important fundamental and applied questions in ecology. However, the data quality from deployed sensors can be difficult and costly to ensure. In this study, we use maintenance records from the 12-year history of Louisiana's Coastwide Reference Monitoring System (CRMS) to assess the relationship between various dimensions of data quality and the frequency of field visits to the sensors. We use hierarchical Bayesian models to estimate the probability of missing data, the probability that a corrective offset of the sensor is required, and the magnitude of required offsets for water elevation and salinity data. We compared these estimates to predetermined risk thresholds to the help identify maintenance schedules that balanced the efficient use of labor resources without sacrificing data quality. We found that the relationship between data quality and increasing maintenance interval varied across metrics. Additionally, for most metrics, the maintenance interval when the metric's credible interval and risk threshold intersected varied throughout the year and with wetland type. These results suggest that complex maintenance schedules, in which field visits vary in frequency throughout the year and with environmental context, are likely to provide the best tradeoff between labor cost and data quality. This analysis demonstrates that quantitative assessment of maintenance records can positively impact the sustainability of long-term data collection projects by helping identify new potential efficiencies in monitoring program management.

Keywords Hydrological recorders · Data integrity · Coastwide Reference Monitoring System · Program management · Hierarchical Bayes

Introduction

The development and deployment of sensor systems and networks in the last two decades have led to a revolution in the spatial scale and temporal resolution of environmental data collection. The richness of this data has allowed ecologists and environmental scientists to measure important processes at relevant scales and allowed for long-term monitoring of environments in ways not previously possible (Collins et al. 2006; Porter et al. 2009; Rode et al. 2016). However, the vast volume and variety of data generated by sensor networks also brings new challenges for quality control

Donald R. Schoolmaster Jr. (✉) · Sarai C. Piazza
U.S. Geological Survey, Wetland and Aquatic Research Center, 700 Cajundome Blvd, Lafayette, LA 70506, USA
e-mail: schoolmasterd@usgs.gov

(Wagner et al. 2006; Campbell et al. 2013; Jones et al. 2018).

Errors in sensor-derived environmental datasets can stem from a number of causes. Some of these causes, such as sensor failure, recorder failure, or disruptive environmental events, leave signatures in the data that can be discovered and corrected by automated methods (Campbell et al. 2013). Other, often more common errors, such as those stemming from sensor fouling and calibration shifts are less easily detected or corrected in post-processing and often require field recalibration (Wagner et al. 2006). As such, for long-term monitoring networks consisting of numerous, spatially disparate sensors, sensor maintenance and field calibration are necessary to ensure data integrity, but can comprise a large proportion of operating budgets.

In order for a sensor-based monitoring program to make informed planning decisions it is necessary to understand the shape of the cost versus data integrity tradeoff. An essential first step in deriving this relationship in determining the relationship between data integrity and maintenance interval. In general, the relationship between data integrity and maintenance is hard to predict de novo since it is likely to depend on a number of factors including the type of sensors used and the characteristics of the local environment where they are installed. However, it is likely that, in some cases, the relationship can be derived from program maintenance records.

In this paper, we present an analysis of the relationship between maintenance interval and a number of different dimensions of data integrity for Louisiana's Coastwide Reference Monitoring System (CRMS). The goal of this analysis is to use the full, 12-year history of CRMS hydrological data to determine the affects of maintenance interval on the probability of missing records and data offset, which we define as difference between the sensor-reported (e.g., water elevation and salinity) and the values that were adjusted relative to an on-site ground truth value. Specifically these analyses focuses on quantifying the relationship between maintenance interval and (1) the probability that an a data error occurred (i.e., missing record occurred or offset required) and (2) the magnitude of the error given that an error was occurred. We demonstrate strategies for how to use this information to select a maintenance interval that maximizes utility on the cost versus data integrity tradeoff manifold.

Methods

Coastwide Reference Monitoring System

History and location

The Federal Coastal Wetlands Planning, Protection, and Restoration Act (CWPPRA) of 1990 was enacted to restore, create, enhance and protect Louisiana's coastal wetlands. Since inception, the CWPPRA program has authorized more than 200 coastal restoration and protection projects. Project types vary by location including river diversions, marsh creation, shoreline protection, vegetative plantings, and hydrologic restoration. As required by CWPPRA, monitoring of each project is mandated throughout the 20-year project life. Prior to 2007, project monitoring was conducted on a project by project basis focusing on paired project and reference sites. Monitoring both project and reference areas provided a means to compare project areas and areas uninfluenced by projects. As the CWPPRA program constructed more than 50 projects, the availability of uninfluenced reference sites, to pair with project sites, was limited. In addition, there was a need to standardize data acquisition, data quality assurance and quality control, and data collection frequency protocols so that the monitoring program could provide data to characterize baseline conditions of Louisiana's extensive coastal wetlands and support landscape scale ecological modeling (Coastal Protection and Restoration Authority 2017).

The Coastwide Reference Monitoring System (CRMS) was designed to provide a long-term reference network that replaced the paired project and reference site monitoring approach. The CRMS network was also designed to monitor the effectiveness of restoration activities at multiple spatial scales, from site to coastwide, because planned restoration and protection activities were intended to influence the entire coastal zone of Louisiana (Steyer et al. 2003). There are approximately 392 CRMS sites representing Fresh, Intermediate, Brackish, and Saline wetland types (Fig. 1).

Data collection methods

Since 2007, each CRMS site has been systematically collecting vegetation, soils, and hydrologic parameters, surface elevation change, accretion, and land to water composition using standardized procedures (Folse et al. 2018). Details of hydrologic sensor specifications, field deployment, sensor servicing, and data processing can be found in the standard operating procedures manual and are summarized below. Hydrologic data (i.e., specific conductance (μS/cm ± 0.5% of reading), salinity (± 0.1 ppt), and water level (± 0.1 ft)) were collected hourly with YSI 600LS continuous data recorder with a vented cable. Field inspections were conducted every 30 to 50 days to download hydrologic data from the continuous data recorder, check for measurement drift, and replace faulty equipment. A portable, hand-held discrete instrument (i.e., YSI 30) was used to verify that the continuous hydrologic data recorder was reading within range and to calculate the measurement drift of the continuous recorder due to biofouling or other causes. At each field inspection, the percent difference between the instantaneous reading for specific conductance on the continuous data recorder and the hand-held instrument was calculated. If the difference exceeded 5%, a data offset (i.e., data adjustment) was applied to the specific conductance data during data processing.The continuous recorder was then cleaned to remove biofouling, and instantaneous readings were recorded again. If the post-cleaning difference exceeded 5%, the continuous recorder was calibrated. Once the equipment was cleaned, and calibrated if necessary, it was redeployed.

To determine if the water level data required a data offset, the water level difference was calculated as the difference between a known mark to sensor distance (as measured by a surveying rod) and the instantaneous water level measurement of the continuous recorder. If the percent difference was 5% or greater during initial readings, then the water level data was offset during data processing. Professional elevation surveys have been conducted such that each CRMS site has a known mark elevation (NAVD 88, ft). To convert field observations of water level (ft) to water

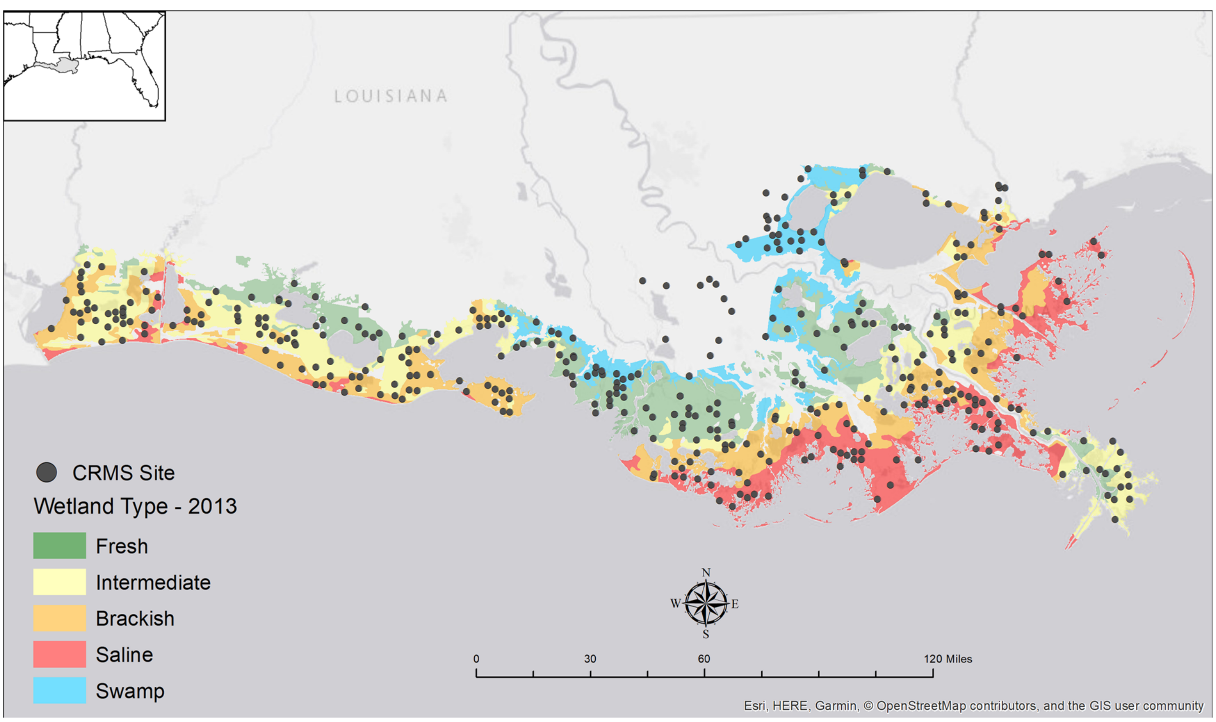

Fig. 1 Map of CRMS sites

surface elevation (NAVD 88, ft) the site-specific mark elevation was applied to the raw water level data.

Data offset methods

When a data offset (i.e., data adjustment) was warranted the specific conductance and water level data were offset during in-office data processing thereby creating a raw and adjusted file for each hydrologic parameter. All raw and adjusted water temperature, specific conductance, salinity, and water elevation data were stored in the Coastal Information Management System (CIMS) database (https://cims.coastal.louisiana.gov/monitoring-data/).

To prevent discontinuous jumps in the archived data, data adjustments were applied gradually over the length the interval as, $\hat{y}_i = y_i + \frac{i}{L}\Delta$, where L is the number of observations in the interval, $i = \{1, 2, \ldots, L\}$, y_i is the raw value of observation i, $\hat{y}_i$ is the adjusted value of observation i, and Δ is offset recorded during the field visit. The transition between data files were checked to ensure that the offset resulted in smooth transitions between previous and subsequent intervals.

Data analysis

Given the set of dates that each sensor was field-inspected, the interval since the last inspection was calculated and the raw water elevation, adjusted water elevation, raw salinity, and adjusted salinity from the given site at the given time (minus 1 h) were retrieved from the database. Cases in which one or both of the water elevation and salinity measurements were missing were removed. This step resulted in 38,617 remaining observations of paired maintenance interval and sensor offsets. The summary for these data are displayed in Table 1. In addition, we analyzed the patterns of missing records.

For these analyses, maintenance interval, i was divided into 5-day increments, $i = ((0–5], (5–10], \ldots, (175–180])$. For each increment, we estimated the mean and standard deviation of offsets and the proportion of missing records. Because required offsets may be either positive or negative, and thus may potentially cancel one another on average these estimates can inform on the tendency for bias of the offset (i.e., whether sensor drift tends to be positive or negative). To gain more useful insight, we further decomposed the data into estimates of the probability that an interval of the given length results in a required offset and the magnitude of the offsets given that an offset was required. To examine the effect of environmental conditions on the maintenance requirements, we repeated the analysis across groups for two partitions of the data, time-of-year and wetland type. For the time-of-year analysis, we divided the offset data into groups based on the season of the field inspection. The groups were as follows: January through March, April through June, July through September, and October through December. To examine variation across wetland types, we divided the water elevation offset data into wetland types based on salinity (Freshwater, Intermediate, Brackish, and Saline) as described by Visser et al. (2002). The patterns of record loss were identical for both water elevation and salinity. As such, salinity was fully analyzed.

Estimation was carried out using Hierarchical Bayesian estimation via JAGS MCMC (Plummer 2016) on the R Platform (R Core Team 2017). In each analysis, the parameters of the distribution of the measured variables at each interval were assumed to be drawn from a distribution with shared hyperparameters. This structure allows for partial pooling across intervals and reduces uncertainty in estimated parameters represented by few samples. The specific models are introduced below and depicted graphically in Fig. 2.

The probability, p_i that an interval of length i would result in an offset or missing record was assumed to be described by a binomial distribution,

$$x_i \sim \text{Bin}\left(n_i, p_i\right), \tag{1}$$

where x_i is the number of times an offset was required (or record was lost) and n_i is the total number of field inspections (records accumulated) after i days. We assumed that the probabilities at each interval were drawn from a common Beta distribution, that is,

$$p_i \sim \text{Beta}\left(\mu, \phi\right). \tag{2}$$

To estimate the magnitude of the offsets given that an offset was required, we created a subset of the data that included only those offsets that were non-zero. We assumed that the absolute value of the adjustments, a_{ji}, were log-normally distributed and fit the parameters of the lognormal distribution,

$$\text{Log}(a_{ji}) \sim N\left(\mu_j, \sigma_j\right). \tag{3}$$

Table 1 Summary of data

	Interval (days)	Elev. Adj. cm)	Sal. Adj. (ppt)	Missing records
Min.	0.1667	−165.45	−10.94	0
Median	33.91	0.00	0.02	0
Mean	39.38	−0.28	0.10	41.14
Max.	270.00	48.16	17.34	5006

We assumed that μ_j and σ_j were drawn from common distributions, i.e.,

$$\mu_j \sim N(\mu_0, \sigma_0),$$
$$\sigma_j^2 \sim \mathrm{Inv}-\mathrm{Gamma}(\alpha, \beta) \tag{4}$$

The mean magnitude of the offsets, given that an offset was required was calculated as $E(P(a_j|a_j \neq 0)) = \exp(\mu_j + \sigma_j^2/2)$. The uncertainty around the means were estimated by MCMC.

The proportion and total number of lost records, y_{ji} at each interval was assumed to be Poisson distributed,

$$y_{ji} \sim \mathrm{Pois}\left(\lambda_{ji}\right). \tag{5}$$

We assumed that λ_{ji} was a function of the total number of records accumulated over the interval N_j and the rate of loss at that interval $1 - e^{r_i}$. Specifically,

$$\lambda_{ji} = N_j e^{r_i}, \tag{6}$$

where N_j is the total number of records in observation j and e^{r_i} is the proportion of intact records characteristic of maintenance interval i. An estimate of the average number of lost records per interval i was calculated as $\overline{y}_i = 24 M_i e_i^r$, where M_i is the upper threshold, in days, represented in maintenance interval i. The uncertainty around these values were estimated by MCMC.

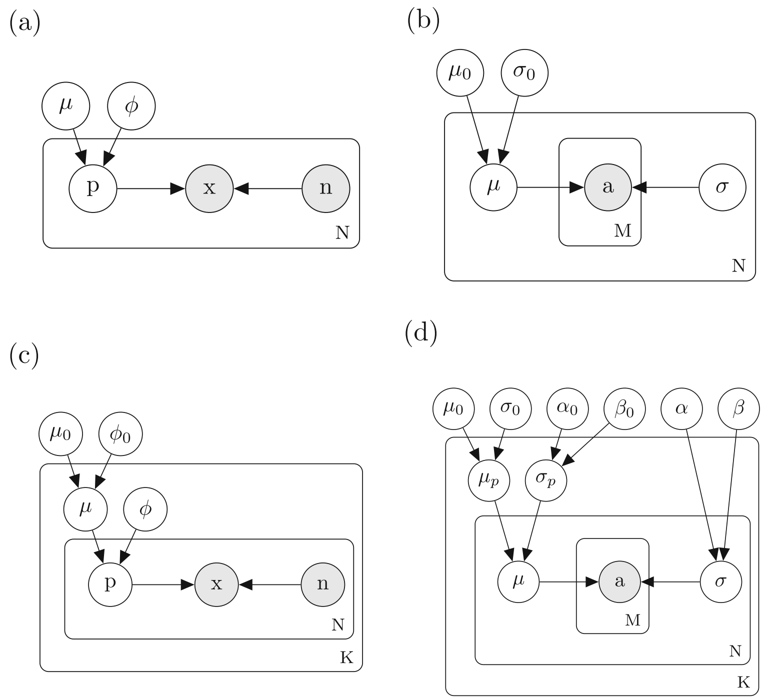

Fig. 2 Directed acyclic graphs (DAG) of hierarchical Bayesian models used for analyses. (**a**) Probability of required offset. (**b**) Magnitude of offset given one is required. (**c**) Probability of required offset for data partitioned by season or wetland type. (**d**) Magnitude of offset given one is required for data partitioned by season or wetland type. Shading indicates observed nodes; unshaded indicates latent nodes. Plates indicate indices and are labeled by the number of elements in the index. For example, plate M in (**b**) and (**d**) indicates observations a_{jik} where $j = 1 \ldots M$ for interval $i = 1 \ldots N$ for data partition $k = 1 \ldots K$. The DAG for the missing records model has the same structure as shown in (**a**), where $x = y$ and $p = r$

For data partitioned by season or wetland type, the hierarchical models used were similar to those described above, but include an extra level in the hierarchy. Specifically, the probability, p_{ij} that an interval of length i in partition j would result in an offset was assumed to be described by a binomial distribution,

$$x_{ij} \sim \text{Bin}\left(n_{ij}, p_{ij}\right), \tag{7}$$

where x_{ij} is the number of times an offset was required (or record was lost) and n_{ij} is the total number of field inspections (records accumulated) after i days. We assumed that the probabilities at each interval were drawn from Beta distribution with partition-dependent parameters, that is,

$$p_{ij} \sim \text{Beta}\left(\mu_j, \phi_j\right). \tag{8}$$

Finally, we assumed that μ_j were drawn from a common Beta distribution,

$$\mu_j \sim \text{Beta}\left(\mu_0, \phi_0\right). \tag{9}$$

Note that in these analyses, the credible intervals calculated represent the precision with which the parameter is estimated, not the deviation of the assumed distributions. We chose this measure of uncertainty to reflect the different amount of information we have for each interval. All MCMCs were run with 4 chains, burned in for 1000 steps, and run for an additional 4000 steps. Convergence was checked using the gelman-diag() function of the coda package (Plummer et al. 2006).

To put the results in a context required for risk-based decision making, we compare the estimated parameters to a priori thresholds of acceptable levels of risk. For this study, conservative thresholds were chosen to balance the efficient use of labor resources without sacrificing data quality. For probability of required offset, we set the threshold at 0.20. For magnitude of offset, given offset is required, we set the threshold at 25 times the error of the sensor. The threshold for lost records was set at 168 records (i.e., 1 week's worth of hourly records). Additional considerations for setting a priori risk thresholds are presented in the discussion.

Results

Across most of the estimates (e.g., Figs. 4, 5, 6, 7, and 8), very short intervals ($<$ 15 days) depart from the patterns represented by the rest of the intervals, often being larger and more variable than the rest of the pattern would suggest. The likely reason is that there were relatively few field visits at such short intervals, and when there were, it was likely because of a suspected problem with the sensors, for example, after a large storm. As a result, when we compare estimates to the risk thresholds, we will ignore the first few intervals.

Each indicator of data quality, elevation offsets, salinity offsets, and lost records showed different patterns of change across maintenance interval (Fig. 3). The average offset for water elevation was effectively constant across intervals of different length and smaller than the accuracy of the sensor, which indicates no tendency for directional bias of sensor errors with increasing maintenance interval. Alternatively, the average offset for salinity showed a pattern of initial positive increase and eventual saturation as maintenance interval increased. In most cases the upper bound of the credible interval was greater than the functional accuracy of the sensor, suggesting a bias toward negative errors. With exception of an unusually high mean rate of lost record production at the shortest maintenance interval (0–5 days), the mean rate of lost records production was low and slightly increasing before becoming much more variable at intervals over 110 days.

The decomposition of the mean offset and lost record rate into the probability of a data quality issue (Fig. 4) (i.e., offset required or record lost) and the magnitude of the event given that one occurred (Fig. 5) uncovers additional differences among the indicators. For water elevation, the probability that an offset was required is low and constant until 100–110 days, which is first maintenance interval for which the credible interval of the probability crosses the risk threshold of 0.2 (Fig. 4a). However, although the probability an offset is required is constant and small for maintenance intervals less than 85–90 days, the magnitude of the adjustment is relatively large, around 5 cm (Fig. 5a). The credible interval of the mean magnitude of the offset crosses the risk threshold for the first time at 65–70 days, although the expected value is never greater than the threshold.

Unlike the pattern for water elevation, after the initial few intervals the probability of a required salinity offset steadily increases with increasing maintenance interval until the credible interval crosses the risk

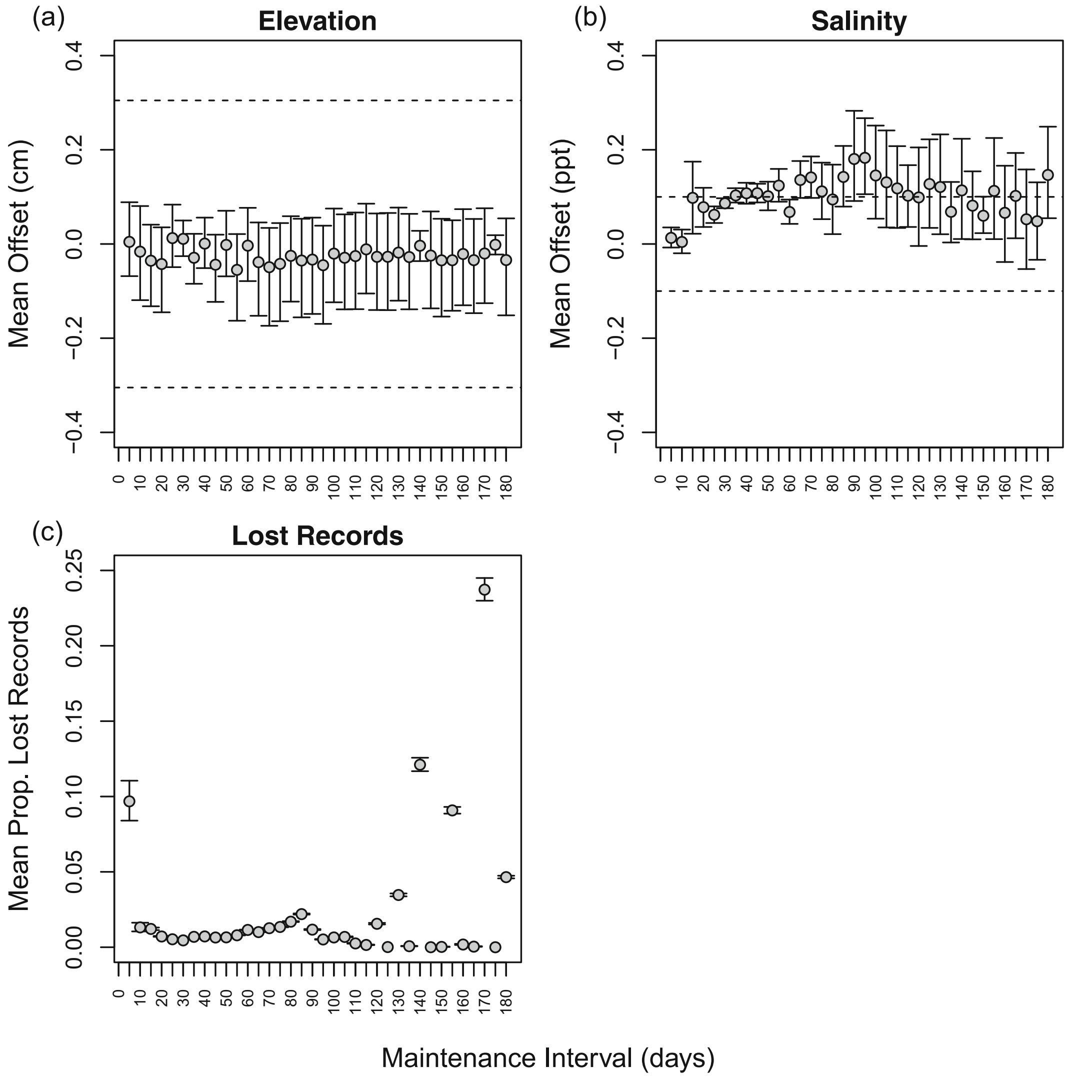

Fig. 3 Mean offset for set of all field inspections for (**a**) water elevation and (**b**) salinity as a function of maintenance interval. (**c**) Mean proportion of lost records as a function of maintenance interval. The dashed lines indicate the working precision of the sensor

threshold at an interval of 85–90 days (Fig. 4b). The range of the credible intervals continues to increase beyond 90–95 days. The mean magnitude of salinity offsets was around 1 ppt and independent of maintenance interval (Fig. 5b). The upper bound of credible interval for mean magnitude of salinity offsets was less than the predetermined acceptable risk threshold across all maintenance intervals.

The probability of at least one lost record was higher than the probability of required offsets for water elevation and salinity across all maintenance intervals (Fig. 4c). Other than the initial few intervals, the probability of lost records was below the acceptable risk threshold but steadily increasing with increasing maintenance interval until 65–70, which is the first interval in which the credible interval of the probability includes the risk threshold. After maintenance intervals of 75–80 days the expected value of the probability of lost records is usually greater than the risk threshold. The mean number of lost records, given that at least one was lost (Fig. 5c), followed a similar pattern to the mean proportion of lost records (Fig. 3c). The mean number of lost records remains below the acceptable risk threshold of 168

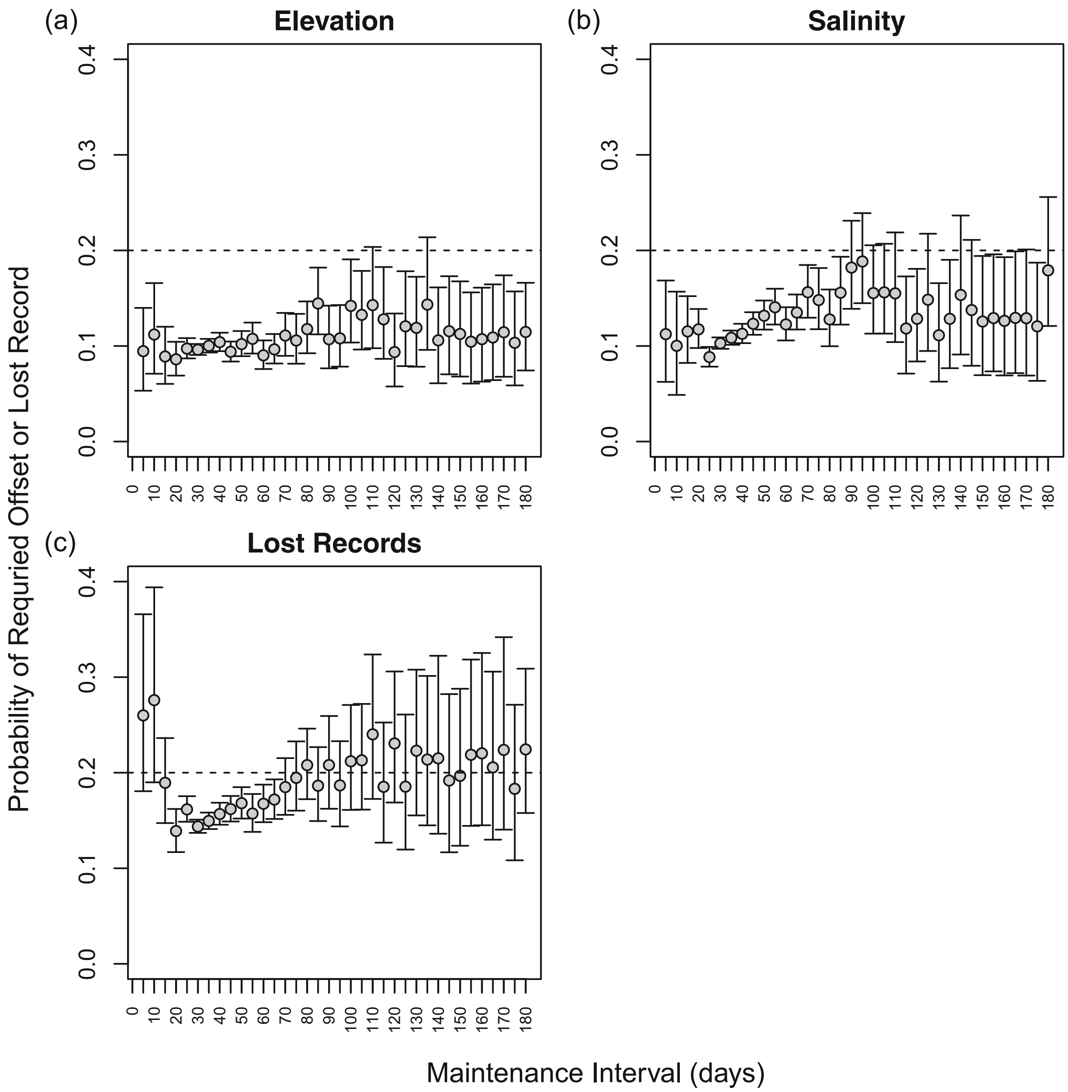

Fig. 4 Probability that a (**a**) water elevation offset was required, (**b**) salinity offset was required, or (**c**) at least one record was lost as a function of maintenance interval. Dashed lines indicate predetermined acceptable risk thresholds

records until 80–85 days. The mean number of records lost becomes much more variable for maintenance intervals beyond 115–120 days.

Variation across wetland type Across wetland types, the probability an elevation offset was required and the mean magnitude of offsets given one was required follow generally similar patterns with increasing maintenance interval. However, the maintenance interval for which the upper bound of probability of required offset included the acceptable risk threshold varied by wetland type (Fig. 6). Brackish marshes were the earliest to have an upper bound of estimated probability be larger than the risk threshold at 80–85 days (Fig. 6c). For Intermediate marshes, it occurred at 105–110 days (Fig. 6b). For Fresh and Saline marshes, the upper bound of estimated probability of a required offset was less than the acceptable risk threshold for all maintenance intervals (Fig. 6a, d).

The mean magnitude of elevation offsets also varied among wetland types. In Fresh marshes, the upper bound of the mean offset was greater than the risk threshold at most maintenance intervals and the expected value of the mean magnitude was greater

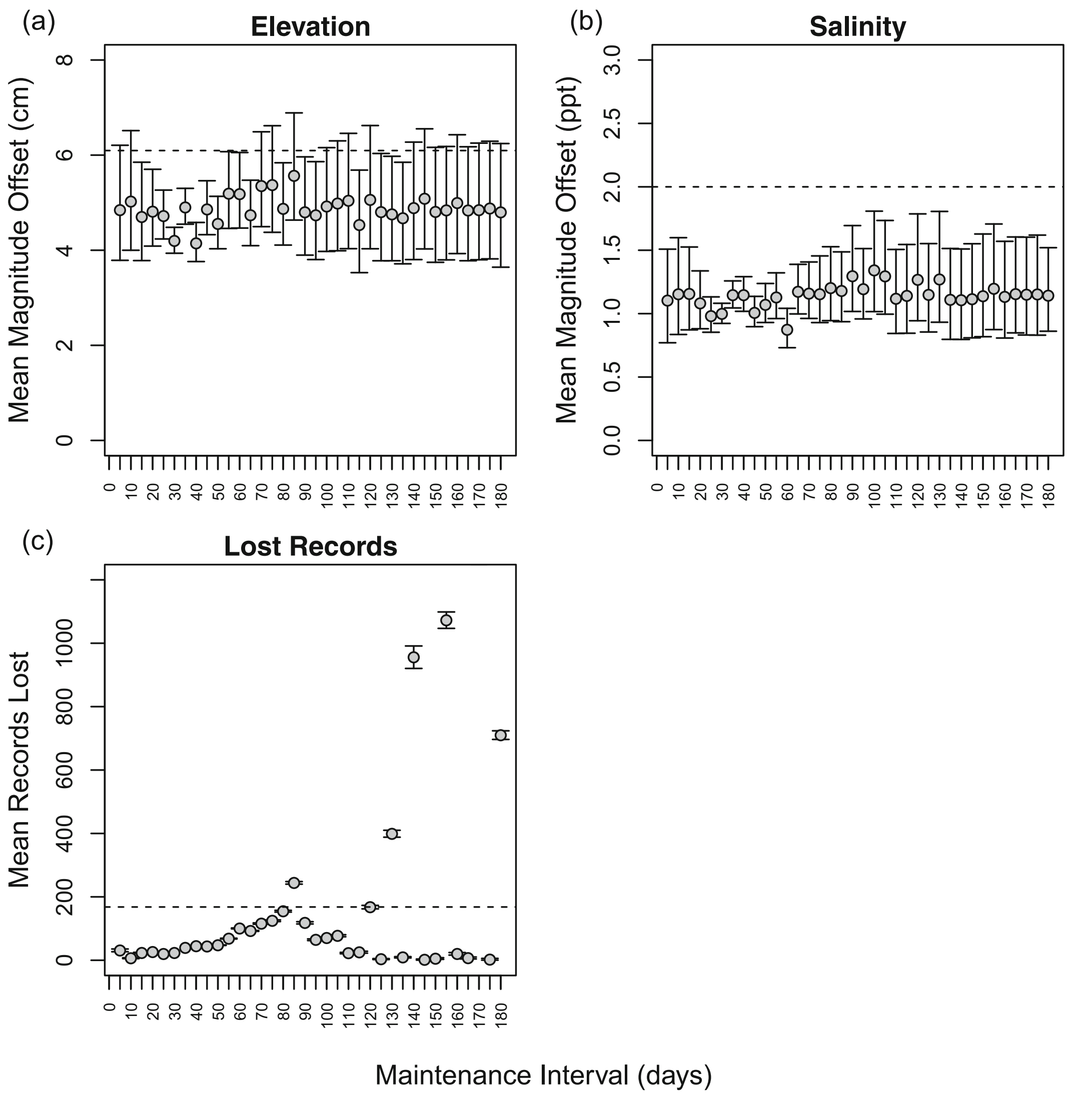

Fig. 5 Mean magnitude of (**a**) water elevation offset given an offset was required, (b) salinity offset given an offset, and (**c**) Number of lost records, given that at least one record was lost as a function of maintenance interval. Dashed lines indicate predetermined acceptable risk thresholds

than the risk threshold for some intervals longer than 30–35 days (Fig. 7a). For Brackish and Saline marshes, expected value of the mean offset magnitude was greater than the risk threshold for a few intervals longer than 70–75 days (Fig. 7c, d). For Intermediate marshes, the upper bound of the mean offset was greater than the risk threshold at most maintenance intervals longer than 60–65 days (Fig. 7b), but the expected values of mean offset magnitude was less than the risk threshold over all maintenance intervals (Fig. 7b).

Variation across seasons There was substantial variation across seasons in the probability of a required elevation offset (Fig. 8). For field visits in January through March, the upper bound of the credible interval was greater than the risk threshold for most intervals longer than 60–65 days (Fig. 8a). Across all the other seasons, there was only one other estimate for which the upper bound of the credible interval was greater than the risk threshold, 95–100 days for field inspections from October through December (Fig. 8c). The mean magnitude of elevation offsets showed a

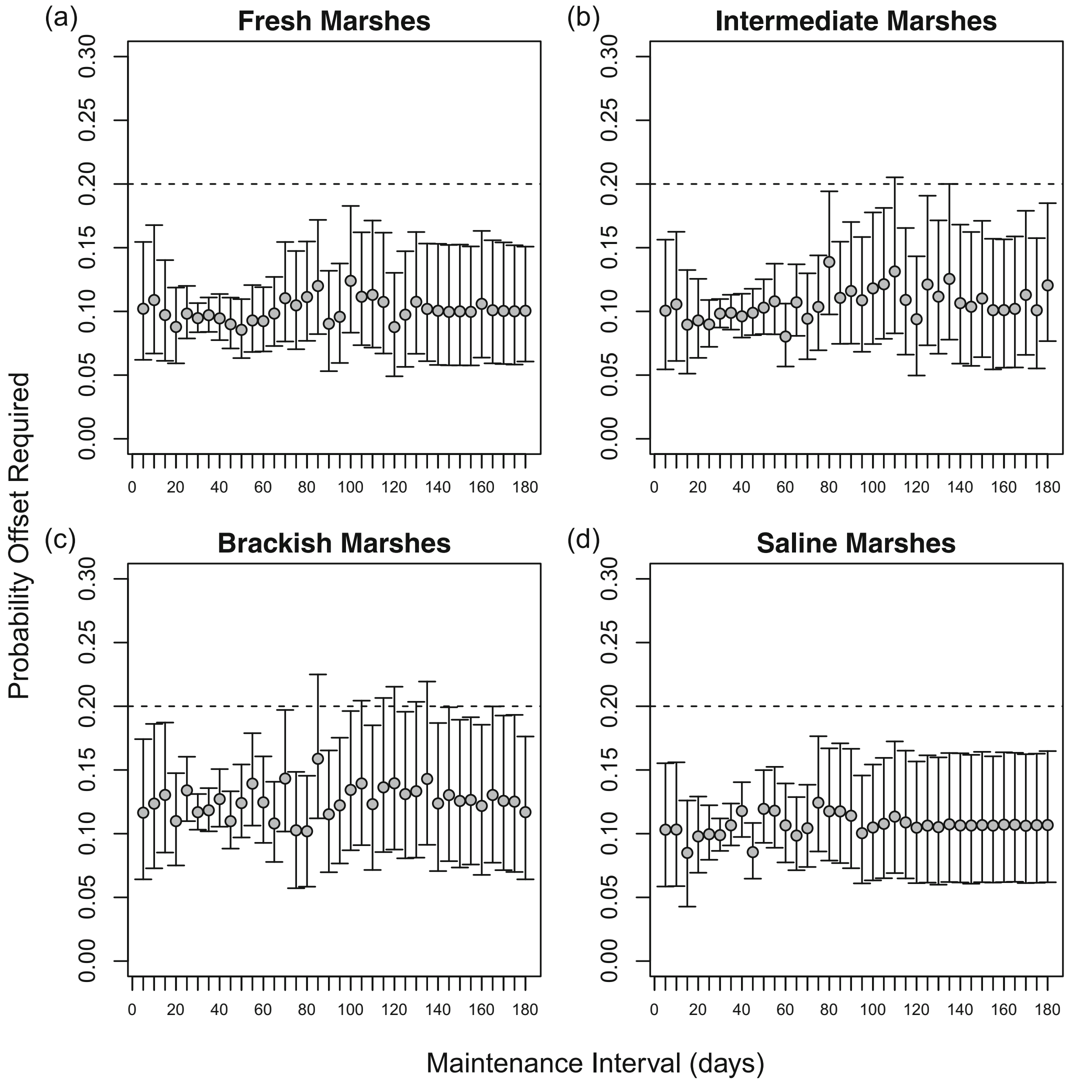

Fig. 6 Probability that a water elevation offset was required as a function of maintenance interval for different wetland types: (**a**) Freshwater marshes (salinity, 0 ppt), (**b**) Intermediate marshes (salinity, 0–5 ppt), (**c**) Brackish marshes (salinity, 5–15 ppt), (**d**) Saline marshes (salinity, $\geq$ 15 ppt). Dashed lines indicate predetermined acceptable risk thresholds

biennial structure, with similar patterns from April through September (Fig. 9b, c) and a different pattern from October through March (Fig. 9a, d). In the first case, the upper bound of the credible interval of offset magnitude was greater than the risk threshold for most maintenance intervals, while in the later, with exception of the first few intervals, the upper bound of the credible interval of offset magnitude was less than the risk threshold for maintenance intervals below 40–45 days.

Discussion

Understanding the relationship between maintenance interval and data quality from sensor-based monitoring networks is crucial for planning and sustainability of monitoring programs. Much of the current discussion around data quality of sensor networks assumes wireless sensors in which the data can be monitored in real-time (e.g., Campbell et al. 2013; Horsburg et al. 2015). When the sensor data is not streamed

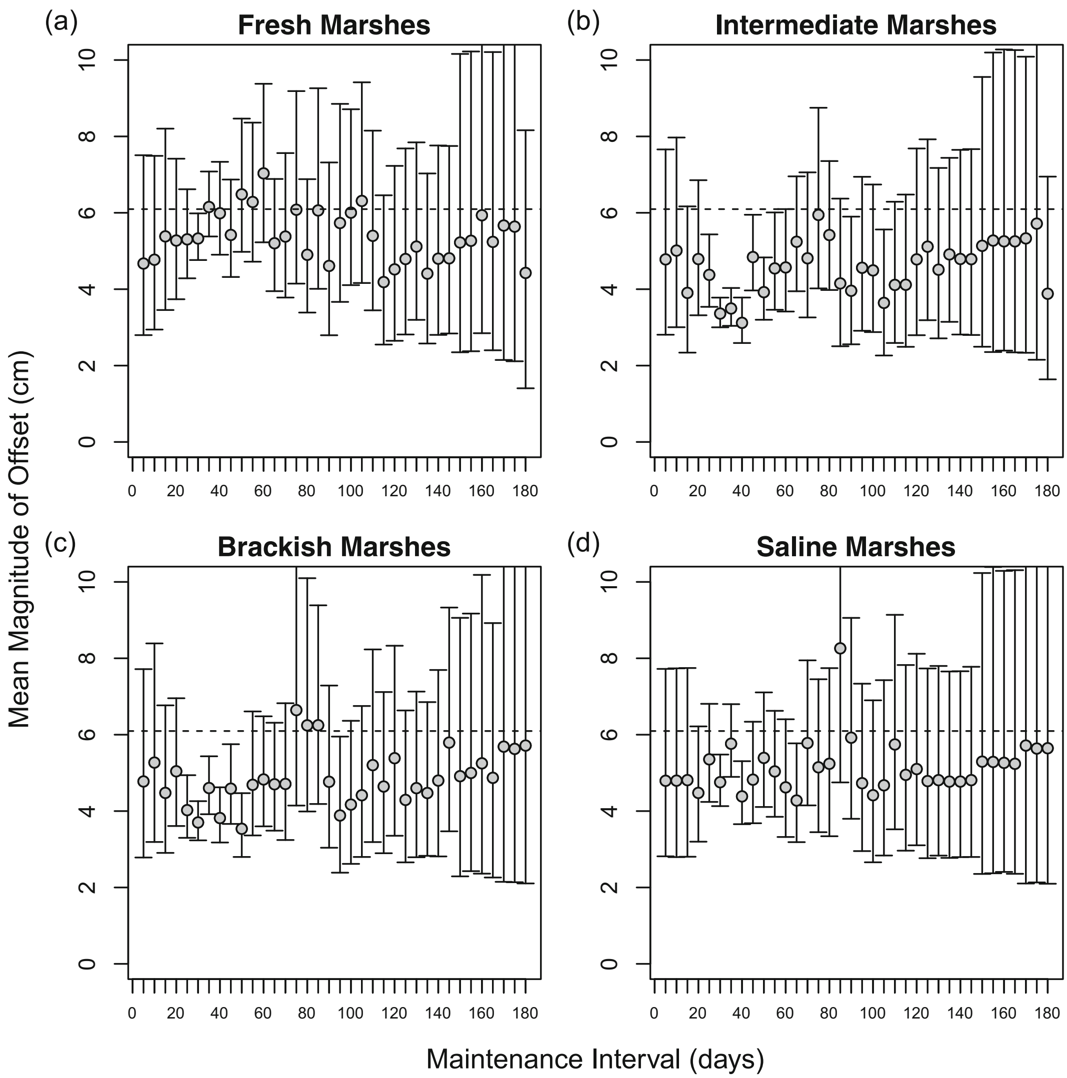

Fig. 7 Mean magnitude of water elevation offset given an offset was required as a function of maintenance interval for different wetland types: (**a**) Freshwater marshes (salinity, 0 ppt), (**b**) Intermediate marshes (salinity, 0–5 ppt), (**c**) Brackish marshes (salinity, 5–15 ppt), (**d**) Saline marshes (salinity, ≥ 15 ppt). Dashed lines indicate predetermined acceptable risk thresholds

to a central location, data quality assurance becomes a function of how frequently the sensors are visited. Using data from the CRMS database, we have shown how maintenance records from field visits and sensor data can be combined to quantify the relationship between maintenance interval and data quality. In addition, we have shown how to combine the derived relationships with a priori risk thresholds to facilitate informed planning decisions with respect to maintenance scheduling.

For the CRMS data, we found that the patterns of the relationships between data quality and maintenance interval varied both among sensor types (Figs. 3, 4, and 5) and within sensors among environmental contexts (Figs. 6 and 7) and across seasons (Figs. 8 and 9). For example, we found that the errors for salinity sensors tended to be biased downward (i.e., indicated water is fresher than it actually is), and reach the risk threshold for probability of required offset sooner (85–90 days) than water elevation sensors

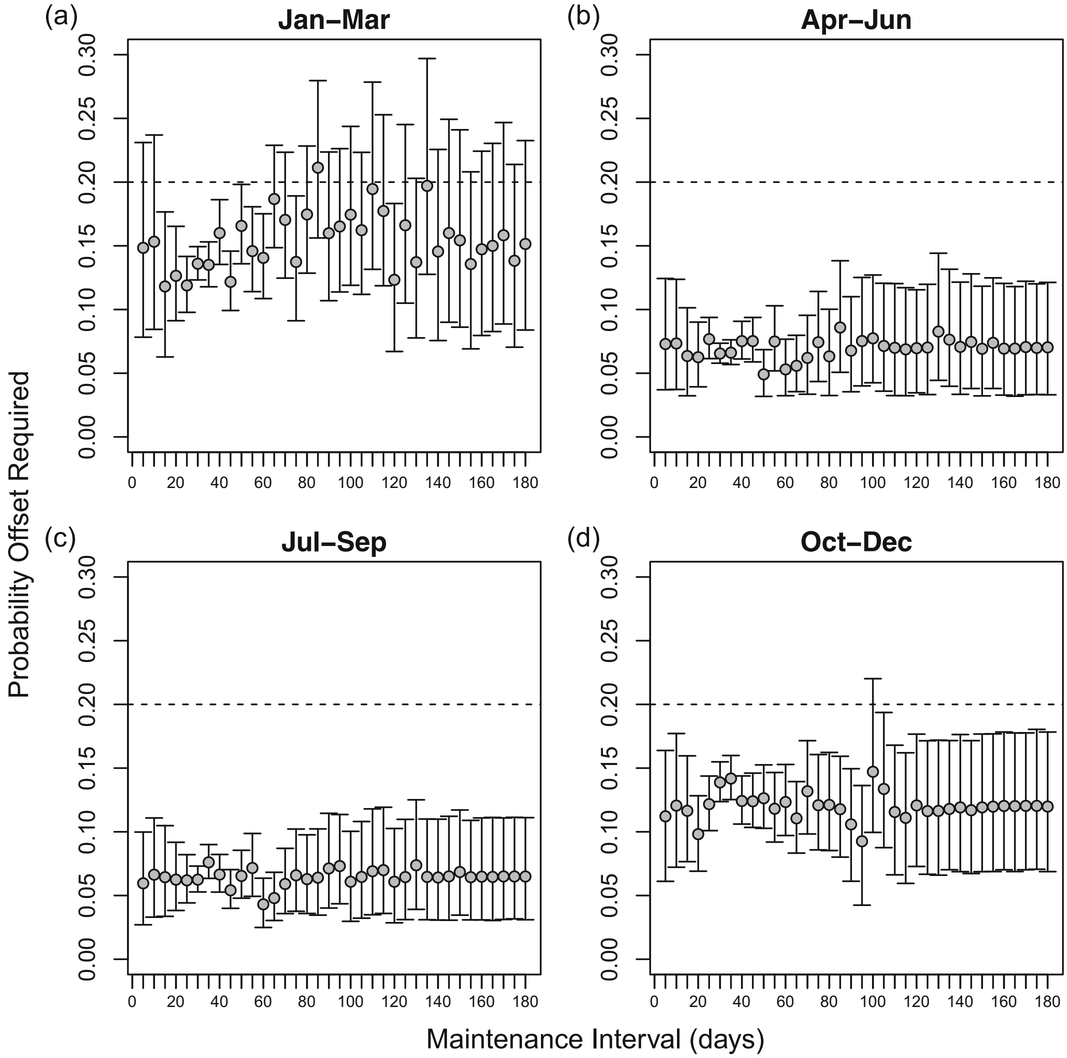

Fig. 8 Probability that a water elevation offset was required as a function of maintenance interval for different seasons: (**a**) January–March, (**b**) April–June, (**c**) July–September, (**d**) October–December. Dashed lines indicate predetermined acceptable risk thresholds

(105–110 days). However, the mean magnitude of water elevation offsets reached the risk threshold after 65–70 days (Fig. 5a), while the salinity offsets never reached the threshold (Fig. 5b). How this information is used to determine optimal maintenance scheduling will depend on the tolerance to each type of risk.

Partitioning the data by wetland type and season suggests that a more complex, finely targeted maintenance schedule could be employed to minimize the data-risk per cost relationship. Our analyses revealed that larger magnitude errors are more likely in Fresh and Saline marshes than in Intermediate and Brackish. This suggests that a maintenance strategy that visits Fresh and Saline wetland types frequently could be a more efficient use of maintenance budget resources than visiting each wetland type on the same interval. Likewise, we also found that required offsets were much more likely in the first quarter of the year (January–March) than in any other, suggesting a maintenance strategy that included more frequent visits during the first quarter of the year.

In this analysis, for illustrative purposes, we used simple, uniform criteria (i.e., a multiple of sensor sensitivity) to set the risk thresholds. In most cases, risk thresholds should be based on specific ecological criteria and monitoring program priorities. For example,

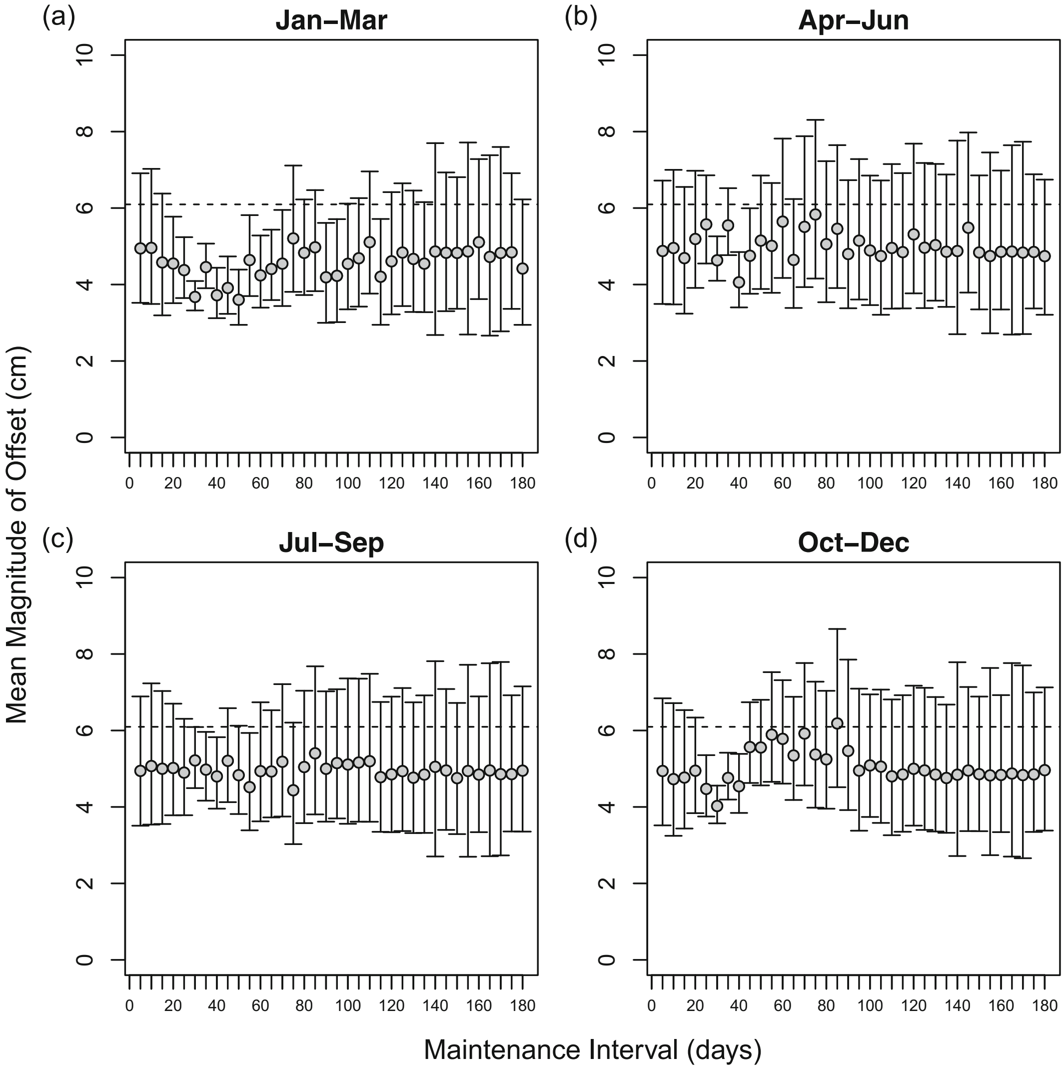

Fig. 9 Mean magnitude of water elevation offset given an offset was required as a function of maintenance interval for different seasons: (**a**) January–March, (**b**) April–June, (**c**) July–September, (**d**) October–December. Dashed lines indicate predetermined acceptable risk thresholds

because of the differential salinity sensitivity of the species in different wetland types, the salinity risk threshold for Freshwater wetlands should be set much lower (e.g., $<$ 1 ppt) than that for Saline wetlands. This is because an increase of 1 ppt salinity in a freshwater system would imply impending ecological shifts, while in a Saline marsh a 1 ppt salinity difference would not. Since there are many potential ways to set risk thresholds, whatever the criteria, it is important for them to be set and clearly justified prior to analysis to prevent bias.

Finally, while this framework was developed from data derived from a mature network, the insights and methods are also relevant to newly deployed networks. Setting acceptable risk thresholds for monitoring targets does not require data and is best done during initial stages of program planning, where the thresholds can also be used to guide decisions as to the required tolerances for the sensors themselves. In addition, because the framework is based on Bayesian inference, estimates may be initially set by expert opinion via prior distributions, and are easily updated as new information arrives through the process of Bayesian updating (Dorazio and Johnson 2003).

Conclusions

As environmental sensor-based monitoring networks mature, the focus must shift from details about deployment to those that allow efficiency and sustainability. These issues can be partially addressed from the data generated by the monitoring program; from sensors themselves (Bandari et al. 1221); and from ancillary sources such as maintenance records. Often, the greatest gains in cost efficiency will come from those elements of the program that are most human-labor intensive, such as maintenance-based field visits. Following a framework, such as the one presented here, that allows quantification of the maintenance cost vs data quality tradeoff has the potential to inform decision making to increase the fiscal sustainability of many sensor-based monitoring networks.

Funding information This work was supported by the CRMS Program, which is administered by CPRA and USGS and is funded through Coastal Wetlands Planning Protection and Restoration Act, the State of Louisiana, with federal resource agencies including the USFWS, NMFS, NRCS, EPA, and USACOE.

Compliance with ethical standards

Disclaimer Any use of trade, firm or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Bandari, S., Bergmann, N., Jurdak, R., Kusy, B. (1221). Time series data analysis of wireless sensor network measurements of temperature. *Sensors*, 17.

Campbell, J.L., Rustad, L.E., Porter, J.H., Taylor, J.R., Dereszynski, E.W., Shanley, J.B., Gries, C., Henshaw, D.L., Martin, M.E., Sheldon, W.M., Boorse, E.R. (2013). Quantity is nothing without quality: automated QA/QC for streaming environmental sensor data. *Bioscience*, *63*(7), 574–585.

Collins, S.L., Bettencourt, L.M.A., Hagberg, A., Brown, R.F., Moore, D.I., Bonito, G.D. (2006). New opportunities in ecological sensing using wireless sensor networks. *Frontiers in Ecology and the Environment*, *4*, 402–407.

Coastal Protection and Restoration Authority. (2017). *Louisiana's comprehensive master plan for a sustainable coast*. Baton Rouge: Coastal Protection and Restoration Authority of Louisiana.

Dorazio, R.M., & Johnson, F.A. (2003). Bayesian inference and decision theory — a framework for decision making in natural resource management. *Ecological Applications*, *13*, 566–563.

Folse, T.M., Sharp, L.A., West, J.L., Hymel, M.K., Troutman, J.P., McGinnis, T.E., Weifenbach, D., Boshart, W.M., Rodrigue, L.B., Richardi, D.C., Wood, W.B., Miller, C.M. (2018). *A Standard operating procedures manual for the coast-wide reference monitoring system-wetlands: methods for site establishment, data collection, and quality assurance/quality control*, (p. 226). Baton Rouge: Louisiana Coastal Protection and Restoration Authority.

Horsburg, J.S., Reeder, S.L., Spackman Jones, A., Meline, J. (2015). Open source software for visualization and quality control of continuous hydrologic and water quality sensor data. *Environmental Modelling & Software*, *70*, 32–44.

Jones, A., Horsburg, J.S., Eiriksson, D.P. (2018). Assessing subjectivity in environmental sensor data post processing via a controlled experiment. *Ecological Informatics*, *46*, 86–96.

Plummer, M. (2016). rjags: Bayesian graphical models using MCMC. R package version 4-6.

Plummer, M., Best, N., Cowles, K., Vines, K. (2006). CODA: convergence diagnosis and output analysis for MCMC. *R News*, *6*, 7–11.

Porter, J.H., Nagy, E., Kratz, T.K., Hanson, P., Collins, S.L., Arzberger, P. (2009). New eyes on the world: advanced sensors for ecology, *59*, 385–397.

R Core Team. (2017). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Rode, M., Wade, A.J., Cohen, M.J., Hensley, R.T., Bowes, M.J., Kirchner, J.W., Arhonditsis, G.B., Jordan, P., Kronvang, B., Halliday, S.J., Skeffington, R.A., Rozemeijer, J.C., Aubert, A.H., Rinke, K., Jomaa, S. (2016). Sensors in the stream: the high-frequency wave of the present. *Environmental Science & Technology*, *50*, 10297–10307.

Steyer, G.D., Sasser, C.E., Visser, J.M., Swenson, E.M., Nyman, J.A., Raynie, R.C. (2003). A proposed coast-wide reference monitoring system for evaluation wetland restoration trajectories in Louisiana. *Environmental Monitoring and Assessment*, *81*, 107–117.

Visser, J.M., Sasser, C.E., Chabreck, R.H., Linscombe, R.G. (2002). The impact of a sever drought on the vegetation of a subtropical estuary. *Estuaries*, *25*, 1185–1196.

Wagner, R.J., Boulger, R.W., Oblinger, C.J., Smith, B.A. (2006). Guidlines and standard procedures for continuous water-quality monitors: station operation, record computation, and data reporting. U.S. Geological Survey Techniques and Methods.