Journal of Hydrology 645 (2024) 132145

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Refining submarine groundwater discharge analysis through nonlinear quantile regression of geochemical time series

Benjamin Hagedorn[a,*], Matthew W. Becker[a], Nyssa J. Silbiger[b], Brittney Maine[a], Ellen Justis[a], Danielle M. Barnas[c], Maya Zeff[d]

[a] Department of Earth Science, California State University, Long Beach 90840, USA
[b] Department of Oceanography, University of Hawaii at Manoa, Honolulu 96822 USA
[c] Hawaii Institute of Marine Biology, University of Hawaii at Manoa, Honolulu 96822 USA
[d] Department of Ecology & Evolutionary Biology, University of California, Santa Cruz 95060 USA

ARTICLE INFO

Keywords:
Multi-Objective Optimization
Radon
Tracers
Coastal Hydrology
Signal Processing

ABSTRACT

A reliable quantification of fresh submarine groundwater discharge (SGD) is key to understanding and managing water quality and habitat of coastal ecosystems. However, current radon tracer-based SGD quantifications suffer greatly from uncertainty and subjectivity in tracer signal processing and the indirect estimation of non-SGD related tracer sources and sinks. In this study we provide a new procedure and model SGD rates as upper envelopes in radon and inverted salinity vs. water depth regressions over tidal cycles. We use non-linear quantile regression to model these envelopes and account for signal-lag correction and noise suppression via optimization towards matched radon and inverted salinity envelope data, and a monotonic (i.e., non-oscillating) trend for the water depth. We then quantify the SGD rate based on the temporal change of the modeled radon envelope inventory and define the radon signature of the fresh groundwater endmember (which is critically important for the calculation of volumetric discharge rates) as the zero-salinity intercept in the envelope regression. Our results indicate that the traditional radon mass balance, if applied at a data aggregation interval of < 1 h and if based on a subjective mixing loss estimation, may lead to a substantial overestimate of SGD. This limitation should be carefully considered in the calibration of lager scale, model-based SGD analyses.

1. Introduction

The analysis of fresh submarine groundwater discharge (SGD) – defined broadly as meteoric water that discharges across a coastal aquifer – is crucial to understanding terrestrial influences on marine ecosystems. SGD of solutes from natural sources or anthropogenic pollution can lead to complex biogeochemical responses that can elevate or harm critical ecosystem functions (Alorda-Kleinglass et al., 2021; Richardson et al., 2017; Silbiger et al., 2020). However, despite its importance, a reliable quantification of the fresh SGD flux has been challenging due to a lack of direct, watershed-specific measurements (Schubert et al., 2019; Taniguchi et al., 2019) and its complex response to tides, wave setup, terrestrial hydraulic gradients and aquifer heterogeneity (Geng et al., 2021; Kreyns et al., 2020; Robinson et al., 2018).

Previous experimental work and modeling studies have shown that fresh SGD from a simplified unconfined coastal aquifer occurs along a narrow discharge "tube", located in-between an intertidal upper saline plume (USP) and the lower saltwater wedge. Along the margins of this "tube", density-driven convection from freshwater and seawater mixing will cause saline SGD; as does tide- or wave action-driven seawater discharge at the base of the USP (Fang et al., 2022; Kuan et al., 2019; Robinson et al., 2018; Xin et al., 2015; Yu et al., 2017). This "tube" fresh SGD model is difficult to apply to complex settings such as tropical volcanic islands and fringing coral reefs, where shallow confining aquitards (e.g., the "reef flat plate") and abundant karstic conduits greatly distort the geometries of the USP, saltwater wedge and SGD "tube" (Fig. 1b; Hagedorn et al., 2020; Houben et al., 2018; Kreyns et al., 2020; Michael et al., 2016). One interesting finding is that aquifer heterogeneity appears to widen the freshwater-seawater interface and to move its toe inland (Geng et al., 2021; Fig. 1b).

The key issue in current SGD research is that tracer inventory-based SGD quantifications, which are critically important for a calibration of model-based SGD estimates for any setting (i.e., with or without a fresh SGD "tube" or USP; Fig. 1a, b), are themselves uncertain. Reasons for

---

* Corresponding author.
E-mail addresses: Klaus.Hagedorn@csulb.edu (B. Hagedorn), Matt.Becker@csulb.edu (M.W. Becker).

https://doi.org/10.1016/j.jhydrol.2024.132145
Received 3 April 2024; Received in revised form 9 September 2024; Accepted 22 September 2024

this include (1) buoyancy-driven geochemical stratification of the water column and the unrealistic assumption of a "well mixed" tracer inventory, (2) the difficult-to-document end-member chemistry of groundwater discharging as SGD, (3) the sensitivity of the SGD estimate to the temporal resolution of the monitoring signals, (4) the challenge of "aligning" signals of SGD tracers from different detection systems with different measurement frequencies, response delays and signal resolution, and (5) the uncertainty from non-SGD related tracer sources and sinks, especially the signal dilution by offshore water; a process commonly referred to as the "mixing loss" or F_{mix} (Adyasari et al., 2023; Savatier and Rocha, 2021; Schubert et al., 2022). A specific problem associated with the F_{mix} parameter is that it is not measured, but typically estimated via subjective interpolation and extrapolation of negative excursions in tracer (i.e., radon) inventory changes. This subjectivity, among other factors, translates into drastic uncertainty in the SGD flux estimate.

In this paper, we present a new methodology for fresh SGD quantification that addresses these challenges. We focus specifically on identifying monitoring periods for which the assumption of a well-mixed tracer inventory holds true and for which a monotonic tracer vs. water depth regression from a tidal forcing can be modeled as a non-linear regression envelope. We define this envelope as a quantile regression function between water depth and the upper extreme of radon and inverted salinity distributions. We attribute positive changes in the envelope inventories during the falling tide (when the hydraulic gradient increases) to the flux of SGD (F_{SGD}), and negative changes during the rising tide (when the gradient decreases) to a low frequency mixing loss ($F_{mix\ (low\ freq)}$). We furthermore attribute upper deviations in the measured radon and inverted salinity signals from the envelopes to statistical (i.e., white) noise and lower deviations from the envelopes to the combined effects of white noise and atmospheric degassing (if only affecting the volatile radon tracer) and a *high frequency* mixing component ($F_{mix\ (high\ freq)}$). The latter reflects the inventory loss from the combined effects of longshore currents and wave action on buoyant freshwater at the stationary SGD monitoring point. The key strengths of our method relative to the traditional mass balance are: (1) it is calibrated because it is based on the envelope function match of two independent SGD tracers: radon and salinity, (2) it has a physical basis and that is the non-linear head loss vs discharge relationship of an unconfined sloping aquifer and (3) it is not dependent on any of the other difficult-to-estimate radon sources and sinks of the traditional mass balance model (e.g., F_{atm} and F_{mix}). We test our method on multiple datasets collected from a tropical, high level volcanic island (Mo'orea, French Polynesia) and demonstrate its relevance and utility for SGD quantifications in other settings.

2. Theoretical background

2.1. SGD estimates from radon mass balance – assumptions and conceptual limitations

The traditional approach for settings, where the effects of surface runoff and evaporation are minor, is to estimate the fresh SGD water flux $F_{\underline{SGDgw}}$ (in m/d) at a stationary monitoring point as (Adyasari et al., 2023; Burnett and Dulaiova, 2003):

$$F_{SGDgw} = \frac{F_{SGD}}{C_{gw}}\left(1 - \frac{S_{mon}}{S_{ref}}\right) \tag{1}$$

where F_{SGD} (in dpm/m^2/d) is the total SGD flux (i.e., fresh and saline SGD) determined from a transient radon mass balance, S_{mon} (in psu) is the measured salinity at the subtidal SGD monitoring point, S_{ref} (in psu) is the reference salinity of non-diluted seawater and C_{gw} (in dpm/m^3) is the radon concentration of the groundwater endmember at the SGD exit point. This value is typically approximated from sampling of inland

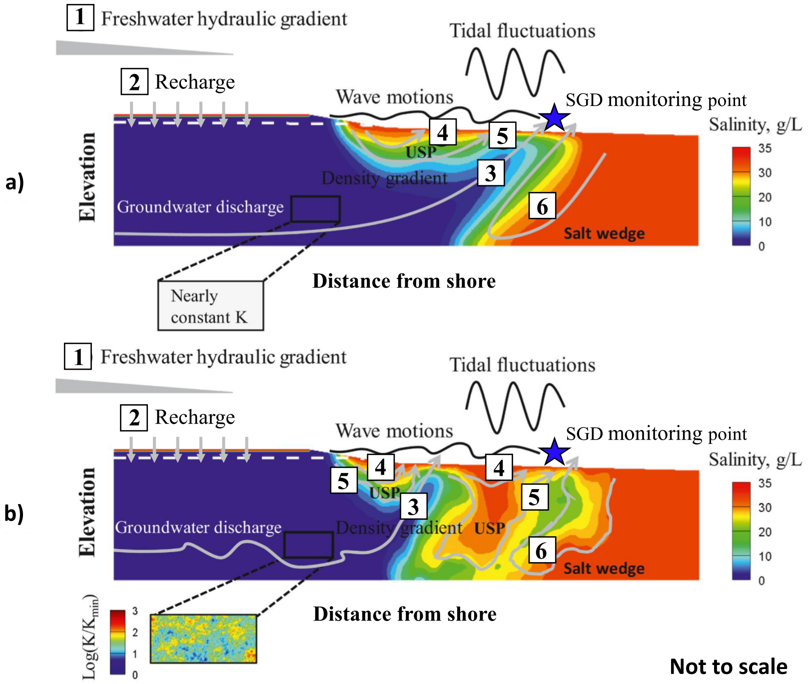


Fig. 1. Conceptual schematics of fresh SGD and saline SGD processes in (a) homogeneous and (b) heterogeneous coastal aquifers (modified from Geng et al., 2021). Everything starts with an inland freshwater hydraulic gradient (1), driven by recharge of fresh meteoric water (2) and causing terrestrial groundwater discharge seaward through the narrow fresh SGD "tube" (3). Processes (4) and (5) indicate saline SGD driven by tides and waves, respectively, in the USP. Process (6) indicates groundwater circulation due to density gradients formed along the dispersed interfaces between fresh water and saltwater. Note the inland shift of the fresh SGD discharge zone in the heterogenous aquifer. Subtidal tracer surveys in (b) (see location of blue star) would yield significantly lower SGD fluxes than in (a). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

groundwater wells that are sometimes far removed from the actual SGD exit point.

The transient radon mass balance applies as:

$$F_{SGD} = F_{net} + F_{mix} \tag{2}$$

with.

$$F_{net} = \frac{dI}{dt} + F_{atm} + F_{dec} + F_{tide\,(ebb)} - F_{tide\,(flood)} - F_{sed} \tag{3}$$

where F_{net} (in dpm/m^2/d) is the net radon flux calculated from measured differences in the radon inventory between successive measurements (dI/dt), and from indirect estimates of the mixing loss (F_{mix}), atmospheric degassing (F_{atm}), radioactive decay (F_{dec}), tidal currents ($F_{tide\ (ebb)}$ and $F_{tide\ (flood)}$) and diffusion from seafloor sediments (F_{sed}). F_{sed} and F_{dec} are typically considered negligible for fringing coral reef settings because of the low-porosity reef caprock and the short duration of tidal cycles, respectively (e.g., Correa et al., 2021; Oehler et al., 2019; Tait et al., 2013, etc.).

The first problem of the method is the assumption of a well-mixed water column and a homogenous radon inventory *I*, which is computed as the product of the measured radon activity and water depth. This assumption only applies to settings where SGD is sampled in

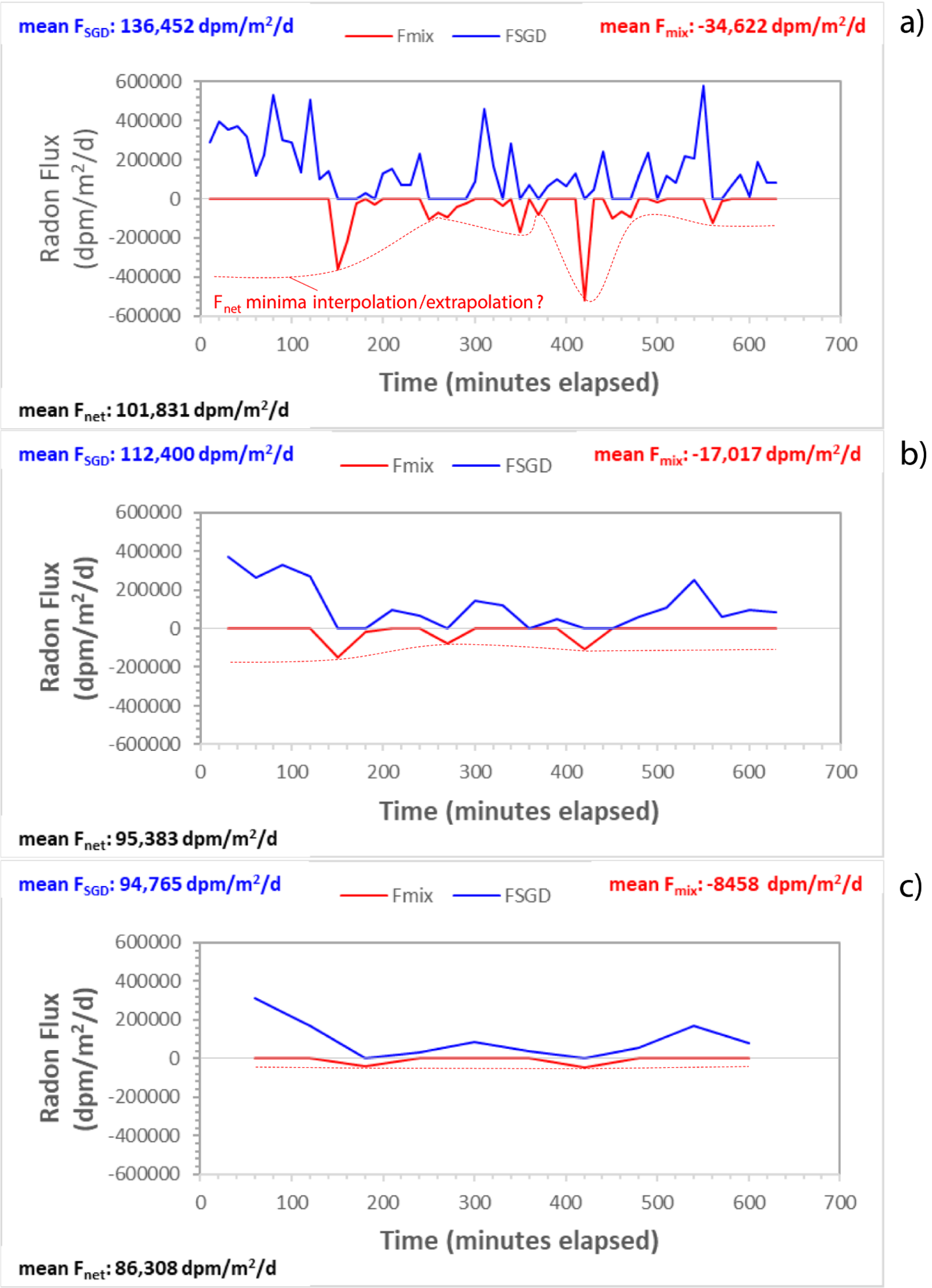


Fig. 2. Example radon mass balance at 10-minute **(a)**, 30-minute **(b)** and 60-minute **(c)** aggregation intervals. Note the attenuating effect of longer intervals on mean radon fluxes. This is especially relevant for F_{mix} because this parameter is typically interpolated and extrapolated over its minima throughout the monitoring period (see red dashed lines). Shorter aggregation intervals yield significantly higher F_{mix} amplitudes for this interpolation/extrapolation. If this interpolation/extrapolation were to be used as a proxy for F_{mix}, the discrepancy in the F_{SGD} estimates for the different aggregation intervals would be significantly greater. See Appendix A for the dataset used to generate this figure. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

shallow water directly from- or slightly above the buoyant freshwater exit point. Away from this exit point, the water column can be extremely stratified, dependent on current direction and magnitude, with freshwater being confined to a shallow surface layer (see, e.g., Johnson et al., 2008; their Fig. 2, or Oehler et al., 2019; their Fig. 3). Radon surveys in a stratified water column will yield overestimates of tracer inventories and unrealistic F_{SGD} rates for equation (1).

The second problem of the method is the assumption that fresh and saline SGD exhibit the same radon activities, but different salinity chemistries so that their relative dominance can be estimated from a simple salinity mass balance (Eq. (1)). This is inconsistent with field studies where saline groundwater sampled within the USP, between the high and low tide marks, yielded significantly lower radon activities than fresh, regional groundwater sampled upgradient of the USP (Colbert et al., 2008a, 2008b; Hagedorn and Tsuda, 2022; Lamontagne et al., 2008). Additionally, in tropical-humid systems, a strong negative linear correlation has been observed between radon activity and salinity values of coastal waters (e.g., Cyronak et al., 2014; Peterson et al., 2009; Su et al., 2014, etc.). This suggests that at well-mapped SGD exit points, the radon contribution from saline SGD throughout tidal cycles is either negligible, or SGD yields salinity vs radon data points that fall on a negative linear fresh groundwater – seawater mixing line. For such a scenario, the $\left(1-\frac{S_{mon}}{S_{ref}}\right)$ term in equation (1) should not be used to specifically differentiate between fresh vs saline SGD, but to distinguish SGD, as a combination of fresh and saline SGD, from seawater.

The third problem of the traditional radon mass balance is the dependency of the model result on the signal processing method. The decision to aggregate, smooth or filter time series data is rarely discussed in the SGD literature, although it has been shown to significantly impact model results in other disciplines (e.g., Alarcon Falconi et al., 2020; Cheng and Adepeju, 2014; Duchesne and Gaillot, 2011; Fefferman et al., 2005; Rostami-Tabar and Mircetic, 2023; Roy, 2020). In relation to SGD, there are two aspects to consider. The first is how the processing method affects the propagation of errors in the F_{net}, F_{SGD} and F_{mix} values. The second is how it causes a delay or *lag* effect for individual signals.

In terms of processing method, most SGD studies follow a simple aggregation approach and "downscale" higher frequency salinity and water depth signals to a longer-interval radon sensor counting cycle that may range from as short as 10 min (e.g., Tait et al., 2013) to up to 2 h (e. g., Burnett and Dulaiova, 2003). A longer aggregation interval will result in a higher signal to noise ratio (SNR), but long intervals may not be able to represent well the temporal SGD dynamics throughout a tidal cycle. A shorter aggregation interval, in turn, translates into a better preservation of higher frequency signals, but also causes lower SNR ratios and thus higher random error signal oscillations. These oscillations have a compounding effect when signals are combined, such as when calculating the radon inventory, I and its change in time dI/dt. The same will be the case for other dependent parameters: F_{atm} (estimated from measured radon activity, temperature, salinity and wind speed-derived solubility and gas transfer coefficients), $F_{tide\ (flood)}$ (estimated as the product of the positive unit change in water depth over the measurement interval times the radon activity of offshore water) and $F_{tide\ (ebb)}$ (estimated as the product of the negative unit change in water depth imes the measured radon concentration in nearshore water). Importantly, the effect of the oscillations will *not* cancel out for F_{net} because the magnitude of $F_{tide\ (ebb)}$ will always be higher than that of $F_{tide\ (flood)}$ due to the former's inclusion of a higher- and time-variant radon signal.

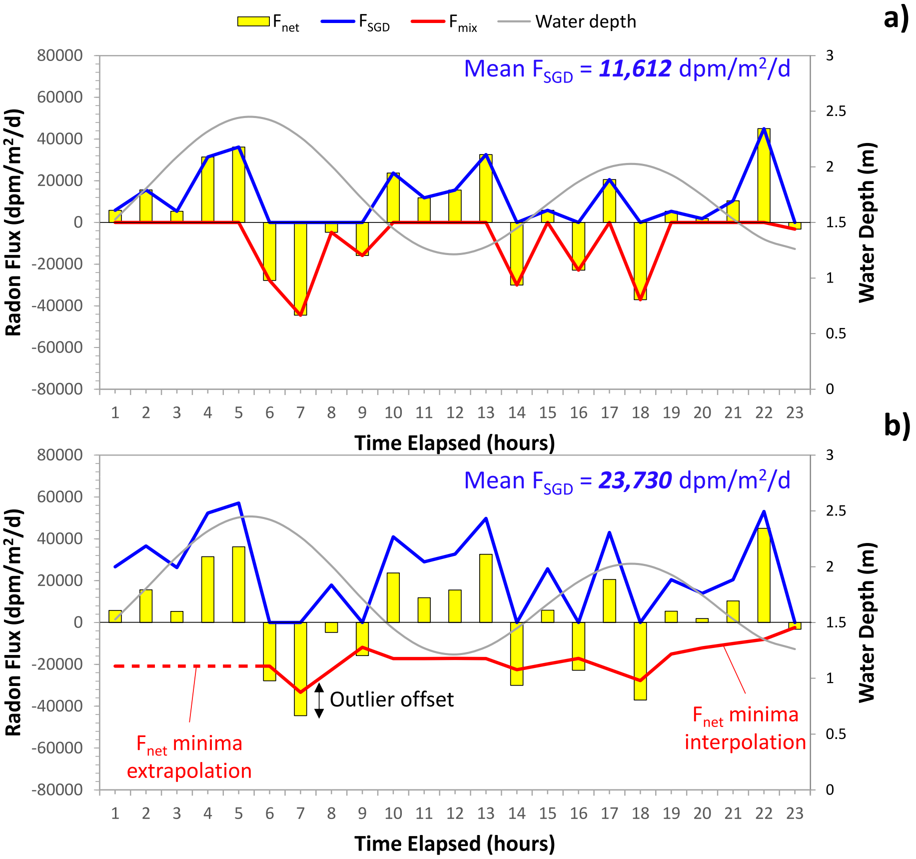


Fig. 3. Case study example (data from Hagedorn and Tsuda, 2022) of radon fluxes from a 24-hour monitoring dataset processed at an 1 h aggregation interval. Because $F_{net} = F_{SGD} - F_{mix}$ and because neither F_{SGD} nor F_{mix} can be measured directly, the latter is commonly estimated based on negative F_{net} (a), or an interpolation/extrapolation of F_{net} minima (b) and the former is estimated as the residual of the equation. Note the drastic effect of the interpolation/extrapolation and its subjectivity on the mean F_{SGD} estimate.

The problem with data aggregation is even more drastic for F_{SGD} and F_{mix} because the distinction between them is essentially categorical: F_{mix} is assumed to equal F_{net} when $F_{SGD} = 0$ and $F_{net} \leq 0$, whereas F_{SGD} is assumed to equal F_{net} when $F_{mix} = 0$ and $F_{net} > 0$. This means that every random error-related positive and negative F_{net} excursion will cause a "false" signal and overestimate of F_{SGD} and F_{mix}, respectively. Fig. 2 a-c illustrates this problem and shows the attenuating effect of a longer aggregation interval (and lower SNR) on the mean F_{net}, F_{SGD} and F_{mix} estimates. In this case, F_{mix} is estimated conservatively to occur only when $F_{net} \leq 0$; i.e., there is no interpolation and extrapolation of F_{net} minima throughout the time series. Importantly, a F_{net} minima interpolation and extrapolation, as commonly done in the literature (e.g., Burnett and Dulaiova, 2003; their Fig. 4), will exacerbate the discrepancy from different aggregation intervals even further (compare red dashed lines in Fig. 2a-c).

The delay or lag effect from signal aggregation is caused by the method's reliance on past values in the average calculation. Because raw water depth, salinity and radon records may exhibit different variances and amplitudes, they will ultimately be subjected to a different delay effect for a specific aggregation interval. The result can be (1) a non-synchronous tracer time series, (2) an erroneous record of the radon inventory and an erroneous radon vs salinity regression. A compounding issue here is that the radon signal exhibits an additional response delay (Petermann and Schubert, 2015; Santos et al., 2012; Stieglitz et al., 2010; Zhao et al., 2022) that varies not only with gas equilibration setup (particularly closed air loop volume and pump flow), but also with equipment wear over time (e.g., clogging of water pump and aeration chamber nozzles by suspended sediment, tubing leaks/kinks, pump flow reductions from declining battery voltage, etc.). While the latter effects may be controlled to some degree, the various response delays of the monitored datasets should still be reevaluated for a specific monitoring campaign. In this study, we highlight the value of a radon vs salinity regression for such a reevaluation.

The fourth and probably most significant problem of the traditional radon mass balance is uncertainty and subjectivity in the indirect estimation of the non-SGD related radon sources and sinks: $F_{tide\ (ebb)}$, $F_{tide\ (flood)}$, F_{atm}, and, most importantly, F_{mix}. It is, for example, not clear how representative the water temperature and wind speed derived F_{atm} loss of radon is at the SGD monitoring point, several cm below the water surface and often 100 s of meters, or even kilometers away from the nearest weather station. Recent research has, in fact, indicated water turbulence (i.e., wave white caps and associated cavitation) as a much more significant driver for F_{atm} than wind speed and water temperature alone (Schubert et al., 2022). Likewise, it is uncertain how reliable the indirect estimate of $F_{tide\ (flood)}$ and $F_{tide\ (ebb)}$ are as radon-enriched or depleted nearshore water may very well be cycled back and forth by consecutive ebb and flood currents. Some studies (e.g., Luo et al., 2020) address this uncertainty using a "return flow factor" that accounts for a specific percentage of the tidal prism that returns from the open sea during a rising tide. The problem here is that reliable measurements of this factor are very difficult to obtain for a specific monitoring site.

The key issue of the F_{mix} parameter is circularity in its estimation because it depends on the same input variables as F_{SGD} (Eqs. (2), (3)). As mentioned before, F_{SGD} is quantified as the positive F_{net}, while F_{mix} is quantified as the negative F_{net}, with most studies interpolating and extrapolating F_{net} minima to occur over the entire monitoring period (Figs. 2, 3; Adyasari et al., 2023; Burnett and Dulaiova, 2003). The purpose here is to allow F_{mix} to occur during low tide and to "increase" the F_{SGD} estimate when F_{net} is positive. Not only is the approach dependent on the magnitude of negative F_{net} excursions and thus on the afore-discussed choice of signal processing method (e.g., short vs long aggregation interval; see Fig. 2), but also is it subjective. There is typically no explanation regarding the interpolation method (e.g., linear vs spline interpolation, etc.). Furthermore, most studies attempt to keep a certain portion of the negative F_{net} below the interpolation line to reduce the influences of negative outliers, but these portions are typically not specified. Lastly, extrapolation beyond any measured F_{net} minima drastically increases uncertainty towards the boundaries. These issues are clearly illustrated in Fig. 3, where Fig. 3a shows a conservative F_{mix} estimation approach without interpolation and extrapolation. In this case, F_{mix} occurs only when $F_{net} < 0$. Fig. 3b shows the more commonly applied F_{mix} interpolation within the measured F_{net} minima and accounting for a negative outlier offset. Also shown is an extrapolation towards the left boundary (see red dashed line). In this example, the subjective F_{mix} interpolation/extrapolation more than doubles the mean F_{SGD} estimate (compare Fig. 3a and b). This raises the question as to why one should not consider the opposite approach and interpolate and extrapolate F_{SGD} through a specific portion of F_{net} maxima and calculate

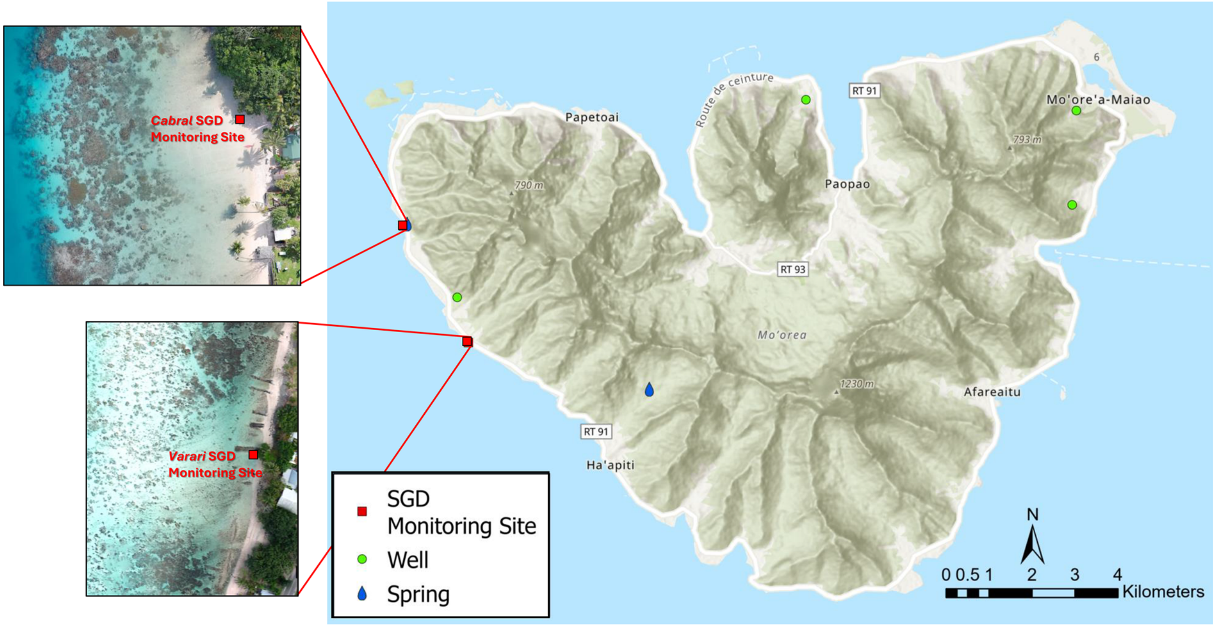


Fig. 4. Location of the study area and sampling sites on the island of Mo'orea in French Polynesia. Inlets depict the Varari and Cabral SGD monitoring sites on the western coast.

F_{mix} as the residual in Eq. (2)? This would allow F_{SGD} to occur irrelevant of tidal stage and even during high tide. From a methodological perspective, such an approach should be considered as "unrealistic" as the commonly applied F_{net} minima interpolation/extrapolation-based F_{mix} estimation.

There is no "right" choice per se, so F_{mix} needs to be constrained independently. Some studies attempt this through measurements (and estimates) of velocities and volumes of moving coastal waters (Oehler et al., 2019; Savatier and Rocha, 2021; Tait et al., 2013; Wolfe et al., 2023), but representative measurements are, again, difficult to obtain at the subtidal SGD monitoring location. An important point to make here is that the mixing loss is not only related to the effects of tide- or wind speed/direction-controlled surface currents on a freshwater plume. Most studies, in fact, sample the water below the water surface very close to the actual freshwater exit point. A more significant factor in this case will be the horizontal back and forth *shearing* effect from wave setup when buoyancy-driven advective flow greatly outweighs diffusion. Because the water column in these settings is mixed only across a very "narrow" band above the discharge point, minor back and forth motions of the buoyant freshwater plume, relative to the stationary SGD monitoring point (i.e., pump inlet), will translate into a highly variable radon inventory and monitoring signal. These variations will be amplified during low tide when the F_{SGD} signal tends to be strongest. The back-and-forth shearing effect causes the homogeneous radon inventory assumption to apply only to a subset of the monitored signal.

2.2. Refining SGD estimates – new approach

To reliably quantify SGD rates, a clear distinction needs to be to be made between the processes attributed to the terrestrial hydraulic gradient and other, higher frequency processes related to other external influences. While F_{SGD} regulates tracer inventory gains (i.e. positive dI/dt; Eq. (3) during the falling tide when the gradient increases, a mixing loss, hereafter referred to as the low frequency mixing loss $F_{mix\ (low\ freq)}$, will always regulate the inventory loss (negative dI/dt) during the rising tides when the gradient decreases. This $F_{mix\ (low\ freq)}$ should not be interpolated or manually "extended" throughout the entire monitoring campaign because its magnitude is controlled by the tidal stage itself. This is different from the "higher" frequency mixing loss, $F_{mix\ (high\ freq)}$, that reflects the more random tracer signal dilutions from longshore currents, wave action, and the overall motion of a buoyant freshwater plume relative to the SGD monitoring point. This mixing loss will happen throughout the monitoring campaign, independent of tidal stage and hydraulic gradient, but quantifying its relative influence, like that of F_{atm} (Eq. (3)), is extremely difficult. Also, inventory gains and losses from F_{SGD} and $F_{mix\ (low\ freq)}$ should not be "normalized" to mean tidal height, as is done with the F_{tide} term in the traditional mass balance, because the tidal height controls the hydraulic gradient and as such the magnitude of F_{SGD} and $F_{mix\ (low\ freq)}$.

Our new approach is to model tide-controlled F_{SGD} and $F_{mix\ (low\ freq)}$ as non-linear envelope functions in regressions of radon and inverted salinity vs. water level data. We refer to the term *envelopes* here as regressions with upper extremes of the response variable distribution and relate these to causal relationships where only a subset of limiting factors applies (Carling et al., 2022). In this case, the envelopes reflect SGD monitoring segments for which the assumption of a homogeneous inventory holds true and for which only the changing magnitudes of F_{SGD} and $F_{mix\ (low\ freq)}$ control the signal. Upper signal deviations from both the radon and inverted salinity envelopes reflect only statistical (i.e., white) noise, whereas lower deviations reflect the combined effects of white noise and F_{atm}, if occurring only for the volatile radon tracer, and $F_{mix\ (high\ freq)}$, if occurring for both tracers, radon and inverted salinity. The key assumptions of our method are: (1) groundwater discharging as SGD exhibits a uniform radon and salinity chemistry, (2) monitoring signals are synchronized; i.e., corrected for statistical or instrument-specific response lags, (3) F_{SGD} and $F_{mix\ (low\ freq)}$ cause mirror-inverted radon and salinity envelope patterns, (4) the capture zone of the SGD monitoring system extends throughout the entire water column, (5) changes in water level are monotonic throughout a tidal segment, thus removing the effect of changing inventory due simply to error in the water level measurements, and (6) the SNRs of monitoring data are maximized; i.e., the F_{SGD} and $F_{mix\ (low\ freq)}$. controlled radon and salinity trends are also monotonic for a given tidal segment. Accordingly, the modeled radon envelope inventory ($I_{envelope}$) in the monitored radon time series can be used to quantify F_{SGD} as:

$$\frac{dI_{envelope}}{dt} = F_{net} = F_{mix\ (low\ frequency)} + F_{SGD} \tag{4}$$

where $F_{mix\ (low\ freq)}$ is considered a loss only when $\frac{dI_{envelope}}{dt}$ is negative. This F_{SGD} can be converted to $F_{\underline{SGDgw}}$ following equation (1). Importantly, however, this F_{SGD} reflects the SGD component distinguished from seawater by concurrently measured radon and salinity data. It, therefore, cannot be further "subdivided" into fresh or saline SGD components, so the $\left(1 - \frac{S_{mon}}{S_{ref}}\right)$ term in equation (1) should be set to one. The C_{gw} term in equation (1) can be estimated as the zero-salinity intercept of the modeled radon and salinity envelope data. The key assumption here is that the radon contribution from saline SGD is negligible and that the fresh groundwater salinity is orders of magnitude lower than the seawater salinity.

2.2.1. Quantile regression

Unlike least squares regression, which estimates the conditional mean of the response variable across values of the predictor variables, quantile regression estimates the conditional median (or other quantiles) of the response variable distribution (Allen et al., 2009; Cade and Noon, 2003; Carling et al., 2021; Koenker, 2017; Lê Cook and Manning, 2013). In situations where there is a weak relationship between the mean of the response variable distribution and the explanatory variables, envelopes (defined as quantile regressions against upper or lower extremes of the response variable distribution) may indicate more useful relationships of how different factors or variables affect different parts of the data (Bergherr, 2018; Cade, 2017).

What makes quantile regression particularly useful for hydrogeologic research is that it extends to non-linear longitudinal data (Chen et al., 2016; Cleophas and Zwinderman, 2021; Geraci, 2019; Hao and Naiman, 2007; Koenker and Park, 1996). This allows it to capture the non-linear, but monotonic head loss vs. discharge relationship of unconfined sloping aquifers (see: Rupp and Selker, 2006 and references therein) and associated tracer concentration changes in seawater affected by groundwater discharge during a falling or rising tide. Another advantage of quantile regression is that it can be used to address the signal noise and associated regression uncertainty of the dependent variable as an exceedance probability through the quantile. This can be determined for different tracers and their detection systems from simple calibration experiments. Similarly, a filter for the explanatory variable (here: water depth) can be applied to minimize the effect of water level signal oscillations from low SNRs on F_{net}. Finally, optimization towards *matched* envelope data of independent tracers – radon and inverted salinity – allows for an objective determination of the appropriate quantiles, filter parameters, tracer signal lags and C_{gw}.

3. Materials and methods

3.1. Study site and datasets

This study focuses on a total of 7 subtidal SGD surveys carried out on the island of Mo'orea, French Polynesia, in August 2021 and March 2022 (Fig. 4). Like other tropical volcanic islands, Mo'orea exhibits a complex hydrogeologic regime with (1) pronounced spatial recharge gradients and (2) transitions from fractured basaltic rock aquifers in the island's interior to a coastal sedimentary wedge of alluvial volcanic sediments, calcareous paleo-reef deposits, and organic debris. This aquifer heterogeneity highlights the challenge of identifying the appropriate C_{gw} value for the F_{SGDgw} calculation (Eq. (1)).

The selection of monitoring sites for this study was based on recent investigations along the western and northern shores of Mo'orea that have revealed low salinities in nearshore waters, indicating potential fresh SGD into the reef (Becker et al., 2023; Hagedorn et al., 2020; Haßler et al., 2019; Knee et al., 2016). The fringing reef along these shorelines is dominated by coral aggregations several meters in diameter, separated by unconsolidated calcareous sand. There is no clear explanation for the documented uneven distribution of fresh SGD on Mo'orea, but results from previous research at other sites (e.g., Cardenas et al., 2010; Correa et al., 2021; Null et al., 2014; Pain et al., 2020) suggest submarine karst-like features, generated by CO_2-rich acidic groundwater, as possible pathways for localized fresh groundwater seepage.

Radon activities and conductivity, temperature and depth (CTD) data were measured at two separate subtidal monitoring sites (hereafter referred to as the *Varari* and *Cabral* sites; Fig. 4) over several tidal cycles. Each monitoring dataset captures peak high and peak low tide changepoints, except for one Cabral falling tide dataset, where technical issues prevented capturing peak low tide. Radon was measured with the RAD7/Aqua monitor (Durridge Company, Inc., Billerica, MA) and CTD was logged concurrently with In Situ Aqua Troll 600 (In-Situ Inc., Fort Collins, CO) and HOBO U24 (Onset Computer Corporation, Bourne, MA) sensors. The detection system was connected to a bilge pump that ran at ~ 5 L/min with its inlet always located about $>$ 10 cm above the mapped SGD seep. This was done to prevent suspended sediment clogging of the pump inlet and the RAD Aqua spray chamber nozzles and to minimize the effect of water stratification that is commonly observed away from the SGD discharge point. As a compromise between allowing high temporal resolution and keeping the RAD7 counting error at a reasonable level, we followed Petermann and Schubert's (2015) approach and selected a 5 min counting cycle for the radon data. CTD readings were aggregated over the same 5 min window. All records were subsequently allocated to the midpoint of each measurement cycle/aggregation window and transformed to a 1 min time-series based on a linear interpolation (see Fig. 5 as an example for a preprocessed dataset). All collected monitoring data of this study are listed in Appendix A.

3.2. Optimization algorithm

An algorithm was coded in R to model the F_{SGD} flux as the upper envelope in radon vs water level regressions combining non-linear quantile regression (nlqr), signal filtering and multi objective optimization. The goal of the algorithm is to find a combination of signal processing parameters that (1) maximizes the correlation between the upper envelopes of the radon vs. water level- and inverted salinity vs. water level regressions and (2) minimizes the effect of error in the water depth signal on F_{net}. Envelope data are then processed according to equations (4) and (1) to quantify $\underline{F_{SGDgw}}$ for the tidal segment. A flowchart detailing the individual processing steps of the algorithm is shown in Fig. 6. The code can be found in Appendix B.

As the first step of the algorithm, we correct the measured radon-in-air data from the RAD7 for the system-specific response delay using the *lag* function of R's *xts* (eXtensible Time Series) package. To account for run- and setup-specific variability, *lag* is set as one adjustable decision variable in the optimization with bounds of ± 60 min. The lagged radon measurements are subsequently converted to radon-in-water activities using water/air partitioning coefficients adapted from Schubert et al (2012), and excess values are calculated by subtracting the dissolved oceanic ^{226}Ra activity (~100 ± 12 dpm/m^3, n = 4). We measured this oceanic activity near our monitoring sites following the methods outlined by Kim et al. (2001) and Wang et al. (2017).

Next, we subdivide monitoring data into either low to high tide (*lh*)- or high to low tide (*hl*) segments to ensure monotonic tracer vs. water level trends for the envelope analysis. This step is necessary because of the hysteretic behavior of water level vs. salinity and water level vs. radon- activity regressions (Fig. 7). Interestingly, there is a contrasting behavior in our monitored data from 2022 (Fig. 7a, b) and 2021 (Fig. 7c, d). Salinity defines an anticlockwise loop from peak high- to peak low tide for the 2022 data and a clockwise sigmoidal loop for the 2021 data. We relate the 2022 trend to a delayed but sudden salinity dilution from freshwater during the falling tide at depths $<$ 0.7 m. The salinity minimum (~31 psu) occurs shortly before peak low tide and values plateau at that minimum well into the rising tide. Then, at depths $>$ 0.7 m, values sharply increase back to the maximum of ~ 37 psu. The 2021 salinity data is different in that there is a much more "immediate" decrease and increase after peak low- and peak high tide, respectively. For radon, all these trends are reversed, i.e., maximum values occur around peak low tide, and vice versa. Similar hysteresis loops have been

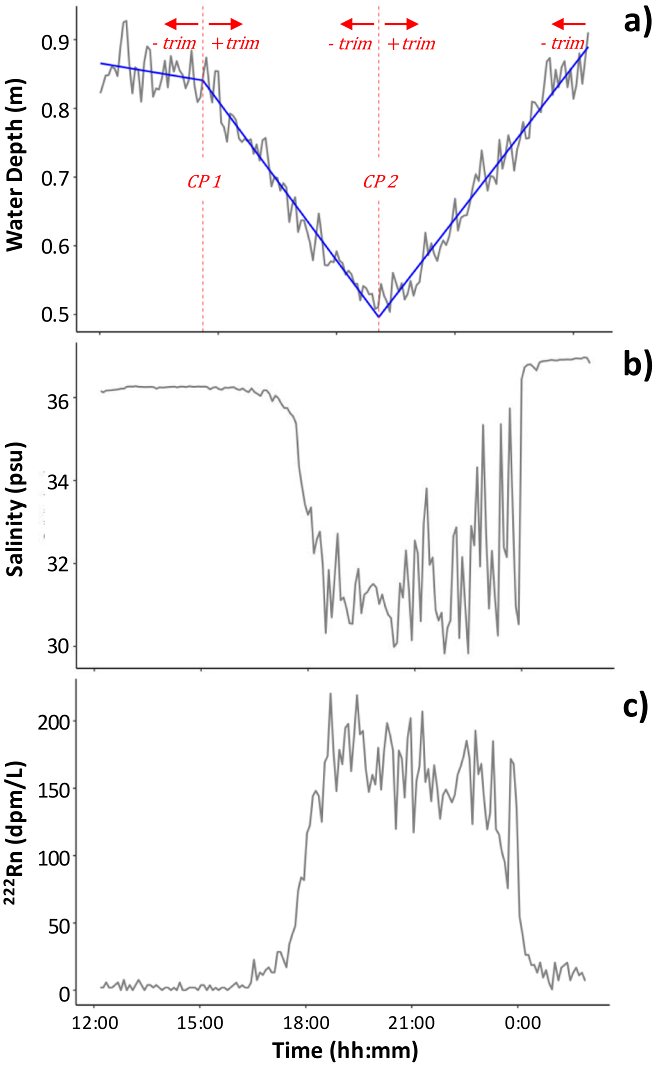
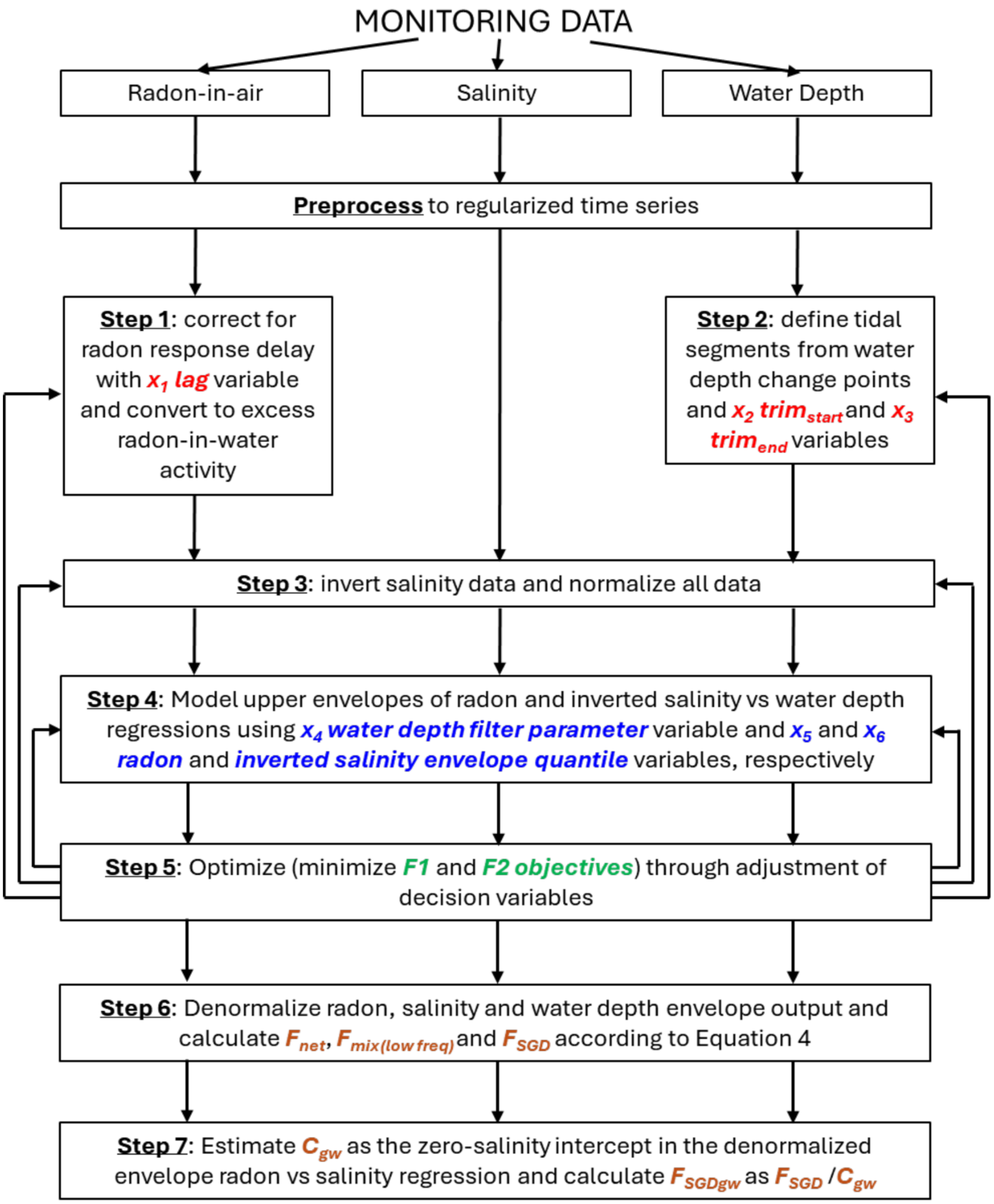


Fig. 5. Unprocessed monitoring data of water level **(a)**, salinity **(b)** and radon activities **(c)** collected at the Varari site from 2022/03/20 to 2022/03/21. Note the constant amplitude in the error related signal oscillations in the water depth data and the changing amplitudes in the radon and salinity signals. High amplitudes in the radon and salinity signal oscillations at low tide reflect complex mixing processes between buoyant freshwater and seawater at the SGD monitoring point. Segmented regression line (blue), change points (red dashed vertical lines) and adjustable change point (CP) trim periods for the optimization algorithm are shown in **(a)**. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



Note: x_1, x_2 and x_3 are **discrete integer decision variables**
x_4, x_5 and x_6 are **continuous decision variables**

Fig. 6. Flowchart showing the coding steps and application of the SGD quantification algorithm.

documented in tracer vs. discharge studies of streams (e.g., Arora et al., 2020; Cartwright et al., 2014; Lloyd et al., 2016; Wymore et al., 2019) and more recently in tidally influenced coastal groundwater (Grande et al., 2023) and signify clearly the complex transitions in the form of the SGD vs hydraulic gradient relationship. Tidal segments for our SGD analysis are defined based on change points (*CPs*) in the water level time series. The *segmented* R package (Muggeo, 2016), that is used for this task, identifies *CPs* iteratively through linearization, maximizing the difference in slope between linear regression lines at the optimized *CP* (see Fig. 5a as an example). To account for uncertainty in the water level data and to allow for a better nlqr model convergence per tidal segment, we apply a trimming decision variable, derived through optimization, at the starting ($trim_{start}$) and the ending ($trim_{end}$) *CP* of each segment as:

$$CP_{high/low\ tide} = CP_{segmented} \pm trim \tag{5}$$

Optimization bounds for the *trim* variable are set to ± 10 min.

In the third step, we change the sign on the salinity dataset (multiplied by negative one) so that all the signals were the same shape. We then range-scale all salinity, radon and water level data as:

$$x, y_{scaled} = \frac{(x, y_{measured} - x, y_{min})}{(x, y_{max} - x, y_{min})} \tag{6}$$

In the fourth step, we model tracer envelopes via nlqr using the *quantreg* R package (Koenker, 2017). These envelopes reflect the F_{SGD}

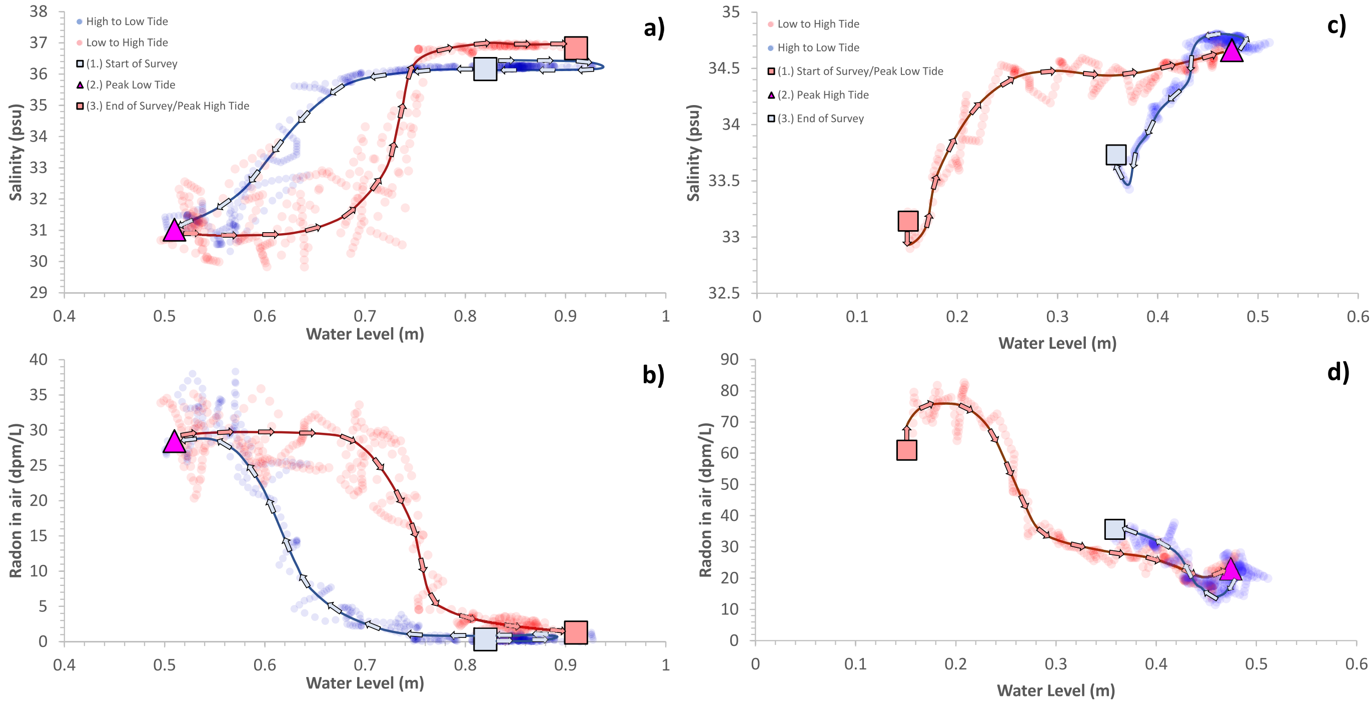


Fig. 7. Hysteresis loops in unprocessed salinity (upper panel) and radon (lower panel) vs water depth regressions from continuous surveys at Varari on 3/21/2022 (**a**, **b**) and Cabral on 08/09/2021 (**c**, **d**). Arrows indicate loop segment pathways from high to low tide and vice versa. Note the opposing pathway directions for different tracers and different sampling campaigns.

and $F_{mix\ (low\ freq)}$ signals over a tidal segment and are expressed as the upper extremes of the radon and inverted salinity distribution against filtered water level data. We define these extremes or "non exceedance probabilities" using quantiles (τ_{Rn} as the radon quantile and τ_S as the inverted salinity quantile) that we determine via optimization. Bounds for τ_S, are 0.9–0.99 based on CTD sensor calibration experiments that indicate a salinity precision error of < 10 %. Bounds for τ_{Rn} are more difficult to assign because of Poisson statistics of radioactive decay and the associated signal resolution dependence on signal strength and counting cycle length (Justus, 2019). Theoretically, τ_{Rn} should change throughout a tidal segment from a low precision (low τ_{Rn}) signal during high tide to a higher precision (high τ_{Rn}) signal during low tide, when the groundwater input of radon is strongest. Noting this uncertainty, we set τ_{Rn} to a broader range of 0.01–0.99.

To maximize the SNR of the water depth data, we apply the Whittaker Henderson (WH) filter, programmed after Eilers (2003) and Weinert (2007), based on a recent filter performance assessment by Schmid et al. (2022). Optimization bounds for the WH filter parameter (λ) are set to 10,000–1,000,000 based on preliminary sample runs that indicate best optimization performance in this parameter range.

For the nlqr function of the envelopes, we select a logistic (sigmoid) curve based on a preliminary analysis of tracer vs. water level hysteresis loop segment shape (see Fig. 7). This curve can be parameterized as:

$$y = y_{\max} + \frac{y_{\min} - y_{\max}}{1 + \left(\frac{x}{c}\right)^{b}} \tag{7}$$

where x and y are the explanatory and dependent variable, respectively, y_{max} and y_{min} are the maximum and minimum asymptotes, c is the point of inflection half-way between y_{max} and y_{min} and b is Hill's slope of the curve at point c. Equation (7) represents a symmetric 4-parameter logistic (4pl) correlation where the asymptotes are defined. This can be simplified to a 3-parameter logistic (3pl) correlation wherein one of the asymptotes is fixed to better handle truncated data (Archontoulis and Miguez, 2015; deAyala, 2021). Both options are considered in this study and the one that provided the best optimization results is used for further processing. To embed the nlqr model in the optimization algorithm, it is programmed to (1) use both the *SSlogis* or *SSfpl* self-starting functions to compute initial estimates of the 3pl and 4pl curve fitting parameters, respectively, (2) apply the non-linear least squares (*nls*) R function to converge on a best fit regression and (3) wrap the code in R's *try* statement to prevent premature model "choking" and ensure continued optimization iterations after *nls* non-convergences (see lines 173 to 183 of the code in Appendix B).

In the fifth step, the two nlqr models per tidal segment are embedded in a multi objective optimization. In this study, we use the Mixed Integer Distributed Ant Colony Optimization (MIDACO) solver (Schlueter et al., 2013; Schlüter et al., 2009) because of its ability to handle both discrete integer- and continuous decision variables and because of its demonstrated success in other computationally expensive applications (e.g., Hagedorn, 2020; Schlueter, 2014; Zobaa, 2019). The first optimization objective, $F1$, minimizes the mean absolute error (MAE; Eq. (8)) between fitted radon and inverted salinity envelope data as:

$$F1 = MAE = \frac{\sum_{i=1}^{n} |Rn_i - S_i|}{n} \tag{8}$$

with $|Rn_i - S_i|$ as the absolute difference between the radon and inverted salinity envelope data. The second objective, $F2$, minimizes the effect of random error in the water depth signal on F_{net} by ensuring a monotonic water depth trend over a specific tidal segment. We accomplish this by maximizing the *Spearman* ρ correlation (Eq. (9)) between water depth and time for each rising or falling tidal segment as:

$$F2 = 1 - \left| 1 - \frac{6\sum d_i^2}{n(n^2 - 1)} \right| \tag{9}$$

where d is the difference between the two ranks of each observation and n is the total number of observations.

In the sixth step, the best fit output for the radon and inverted salinity envelopes are de-normalized and processed according to equation (4) to calculate mean values of F_{net}, $F_{mix\ (low\ freq)}$ and F_{SGD}. The latter is then converted in step 7 to the freshwater flux, F_{SGDgw}, by setting $\left(1 - \frac{S_{mon}}{S_{ref}}\right)$

to one and by determining C_{gw} for each tidal segment via linear regression (i.e., zero-salinity intercept) of the envelope data.

4. Results and discussion

4.1. Quantile regression envelopes

Modeled results for the 7 tidal segments considered in this study are listed in Table 1 and 2. Very low MAE values for objective $F1$ (mean value: 0.052) indicate a very good match between inverted salinity and radon envelopes. That each model converged on a perfect *Spearman* ρ of 1 ($F2 = 0$) renders the second objective as an effective constraint, ensuring a minimized objective 1 at a monotonic water depth. The 3pl model outperformed the 4pl for all but two of the tidal segments; an observation in line with most of the data not being asymptotic with either 0, 0 or 1, 1 (Fig. 8). There is a broad range in the optimized filter parameter λ (34,615 – 832,534; mean value: 251,094) that reflects the different degrees of noise in the different water level and tracer datasets and, consequently, the different degrees of filtering required to limit $F2$ to zero. The encountered λ, τ_{Rn} and τ_S ranges confirm the importance of a campaign-specific signal filtering procedure for a reliable SGD analysis.

Five of the 7 tidal segments revealed lower τ_{Rn} than τ_S values (Table 1) which suggests an overall lower SNR and thus more positive outliers in the radon signal. The discrepancy between τ_{Rn} than τ_S is particularly high for the Varari 03/28/22 *lh* segment (Table 1). While the low τ_{Rn} for this segment (0.699) could be attributed to radon detection outliers related to the often-reported "tailing" or "memory effect" when the radon signal drops (Adyasari et al., 2023; Petermann and Schubert, 2015), it is not clear why the discrepancy is so particularly pronounced in the Varari 03/28/22 *lh* segment, as compared to the other low-to high tide segments. Rather, positive excursions in the inverted salinity signal at low tide (see Fig. 8a, right panel) could be responsible for an exceptionally high τ_S value, and, consequently, a pronounced τ_{Rn} and τ_S discrepancy. The high modeled τ_{Rn} for the Varari 08/04/2021 *lh* and Varari 08/05/2021 *hl* segments (0.941 – 0.988) likely reflects the extreme degassing loss of radon as these datasets were collected during a period of water turbulence with pronounced wave white caps. It appears that the algorithm counteracts the degassing loss by shifting the radon envelope towards a higher quantile. Interestingly, this turbulent water period is not reflected in the wind speed data from nearby (i.e., 2 km north) *Hotel Tipaniers* weather station which yields a lower mean windspeed value for the 08/04/21 and 08/05/21 tidal segments (2.88 m/s) than that for the entire dataset of (4.64 m/s; data from *windguru* online portal). This finding underscores the limitation of indirect degassing loss estimates from wind speed measurements collected at "nearby" weather stations (Schubert et al., 2022, 2019).

The optimization for 3 of the 7 datasets yielded a negative lag (i.e., a backward shift) for the radon time series (Table 1). This is surprising because the response delay from radon detectors typically requires a forward shift. However, it is important to point out here that the lag parameter is optimized for a match between radon and salinity envelope- and not raw regression data (Fig. 9). In situations where the variance in the radon signal is low relative to that of the salinity signal (e.g., Fig. 8a and c), the upper salinity envelope will span a higher water depth range which, because water depth is proportional to time, will cause the salinity envelope to "lag" the radon envelope. The situation for the Varari 08/04/21 *lh* segment is different because the variance in the salinity signal is relatively low (compare mid and right panels of Fig. 8g). In this case, however, the radon signal appears to be significantly affected by the aforementioned "tailing" effect and the negative radon lag is simply required to match the steep 3pl pattern of the salinity envelope.

4.2. Optimized C_{gw} estimates

Regressions of modeled envelopes define strong linear correlations that differ from those of the unprocessed input data (Fig. 9). The difference is particularly drastic for the high-water turbulence datasets, Varari 08/05/21 *hl* and Varari 08/04/21 *lh* (Fig. 9f and g), where extensive radon degassing is evident. Zero-salinity intercepts of all the envelope data reveal a vast radon activity range (Table 2) with mean (441,593 dpm/m^3) and standard deviation (466,965 dpm/m^3) values that are still lower than what we encountered for 4 groundwater and two spring samples from Mo'orea (mean value: 1,104,056 dpm/m^3, standard deviation: 610,089 dpm/m^3; Table 3). The groundwater sample collected closest to our study area (*Coastal Spring*, which is located near the Cabral site) was pumped from a limestone aquifer, whereas all other groundwater samples were collected from basaltic fractured rock aquifers. Importantly, the radon range in groundwater samples not only reflects compositional differences of the aquifers, but also textural factors such as grain shape and size and pore material density (Barillon et al., 2005). These parameters are difficult to account for at a scale that integrates the groundwater flow path form the recharge entry point in the mountainous interior to the SGD exit point along the coast.

The encountered range of envelope zero salinity intercepts reflect these factors. Interestingly, the radon endmembers determined from consecutive segments at the Varari site (03/21/222 *hl* and *lh*) reveal a decrease from 235,095 dpm/m^3 to 168,310 dpm/m^3, that may indicate (1) a changing groundwater radon signal from, e.g., recharge dilution by an antecedent rain event, and/or (2) saline SGD of tidally recirculated seawater affecting the signal of one tidal segment to a greater degree. The similarity of the envelope patterns for radon and inverted salinity for both segments indicates mechanism (1) as the more likely explanation, as a shift towards more saline SGD should coincide with a flattening of the salinity envelope amplitude. Longer-term radon, salinity and groundwater level monitoring are necessary to confirm this. The more

Table 1
nlqr results for high to low tide (*hl*) and low to high tide (*lh*) segments.

Parameter	Varari 2022/03/28 *lh*	Varari 2022/03/21 *hl*	Varari 2022/03/21 *lh*	Varari 2021/08/09 *lh*	Varari 2021/08/09 *hl*	Varari 2021/08/05 *hl*	Varari 2021/08/04 *lh*
Best-fit nlqr function	3pl	3pl	4pl	4pl	3pl	3pl	3pl
Objective Functions							
Objective function F1	0.026	0.003	0.048	0.058	0.189	0.022	0.021
Objective function F2	0	0	0	0	0	0	0
Decision Variables							
τ_{Rn} (−)	0.699	0.960	0.931	0.917	0.628	0.941	0.988
τ_{Sal} (−)	0.925	0.909	0.992	0.989	0.999	0.999	0.904
lag (min)	−31	0	−8	37	1	59	−25
$trim_{start}$ (min)	−10	10	−10	1	−10	−10	6
$trim_{end}$ (min)	10	−10	1	−9	1	3	1
λ (−)	457,117	832,534	34,615	97,684	45,266	244,016	46,426

Table 2
Flux estimates from nlqr-optimized envelope data.

Parameter	Varari 2022/03/28 *lh*	Varari 2022/03/21 *hl*	Varari 2022/03/21 *lh*	Varari 2021/08/09 *lh*	Varari 2021/08/09 *hl*	Varari 2021/08/05 *hl*	Varari 2021/08/04 *lh*
Optimized Results							
mean F_{net} (dpm/m^2/d)	−67,524	91,946	−61,392	12,550	19,307	147,640	−13,015
mean F_{mix} (dpm/m^2/d)	−71,274	−2398	−90,238	−24,223	−243	0	−14,734
mean F_{SGD} (dpm/m^2/d)	3750	94,344	28,846	36,773	19,550	147,640	1719
C_{GW} (dpm/m$_3$)	709,109	235,095	168,310	1,355,500	481,616	132,340	9180
mean $F_{\underline{SGDw}}$ (m/d)	0.005	0.401	0.171	0.030	0.041	1.12	0.187

pronounced endmember decrease modeled for the Cabral 08/09/2021 *lh* to *hl* segments (1,355,500 dpm/m^3 to 481,616 dpm/m^3), however, likely reflects the shortened *hl* dataset, as we were not able to extend our monitoring to peak low tide at < 0.3 m water depth (compare Fig. 8d and e).

Considering all these observations, and the documented groundwater radon variability reported in SGD studies from other apparently monolithologic regimes (e.g., Colbert et al., 2008a; Hagedorn and Tsuda, 2022; Santos et al., 2010; Tait et al., 2013, etc.), we strongly caution against the use of radon data from groundwater well samples as proxies for C_{gw} in Equation (1). Even if the groundwater radon is determined from beach piezometers near an SGD exit point, this groundwater may not be representative for confined groundwater flowing through karstic conduits and discharging as SGD further offshore. As a more reliable approach, we recommend using the zero-salinity intercept of the envelope radon vs. salinity regression as a proxy for C_{gw}.

4.3. Optimized F_{SGD} estimates

Comparing our modeled F_{SGD} time series to those derived from the processed optimization data subjected to the traditional mass balance (Fig. 10) reveals the low pass filter characteristic of the optimization algorithm, where high frequency noise in both the raw tracer and water level time series are clearly suppressed. The mean F_{SGD} estimates for the two successive tidal segment datasets considered in this study (i.e., ~32,809 dpm/m^2/d for Varari 03/21/2022 and 28,161 dpm/m^2/d for Cabral 08/09/2021) fall within the documented range reported for other tropical high-level volcanic islands (Moosdorf et al., 2015).

However, any accuracy assessment for our modeled F_{SGD} rates is complicated by the lack of any measured F_{SGD} calibration data from, e.g., seepage meters and the traditional mass balance's dependence on hard- or even impossible-to-measure input parameters. Fig. 11a-c compares our modeled F_{net}, F_{mix} and F_{SGD} values to those derived from the traditional radon mass balance. In this case, we apply the mass balance to the unprocessed input data (i.e., no lag or response delay correlation) without any F_{net} minima interpolation (i.e., $F_{mix} = F_{net}$ if $F_{net} < 0$). We account for uncertainty only by using both 30- and 60- minute aggregation intervals, but it should be noted that the lowering of the mean F_{mix} through the typically applied F_{net} minima interpolation/extrapolation would increase this uncertainty greatly (Fig. 2). Nevertheless, there are significant positive correlations in each of the regressions ($r^2 > 0.571$, $p < 0.05$) and there is also a tendency for the traditional mass balance, particularly at the shorter aggregation interval, to overpredict F_{net}, F_{mix} and F_{SGD}. The overpredictions are most pronounced for the two successive Cabral 08/09/2021 *lh* and *hl* segments for which the traditional mass balance appears to overpredict F_{atm} due to the high wind speeds at the nearby *Hotel Tipaniers* weather station (Fig. 12). This suggests that these wind speed data are not representative for local conditions during monitoring. Importantly, we were able to match our modeled F_{net}, F_{mix} and F_{SGD} values with the respective mass balance values for all segments through manual adjustments of F_{tide} and F_{atm}. However, we were not able to accomplish this match for scenarios in which (1) data were aggregated to < 30-minute intervals, and (2) F_{mix} was parameterized through an interpolation of F_{net} minima. Our findings thus indicate a tendency of the traditional mass balance to overpredict SGD rates. A subjective F_{net} minima interpolation greatly exacerbates this issue.

4.4. Application of method to other sites

Our method can be applied to any other radon, salinity and water level monitoring dataset. However, only very few studies in the SGD literature publish the raw radon and CTD monitoring datasets needed for such an analysis. An even bigger issue is that processed radon, salinity and depth time series are typically given at a very low temporal resolution (i.e., 30 to 60-minute aggregation intervals). This will cause statistical lag for the different datasets and a low sample size per tidal segment (6 – 12 data points). It will, furthermore, complicate the definition of an envelope line from monitoring segments for which the combined effects of $F_{mix\ (high\ freq)}$ and F_{atm} are, in fact, muted. Collectively, these factors will cause F_{SGD} estimates from quantile regression of low temporal resolution data to be biased low. We thus urge future SGD monitoring studies to report raw radon and CTD monitoring data at aggregation windows of < 10 min.

Nevertheless, simple example applications of our method to 30-minute aggregation interval data from Maui, Hawaii and Dor Beach, Israel (data from Bishop et al., 2017; Weinstein et al., 2007b, 2007a) highlight our method's ability to define quantile regression envelopes to coarser aggregation interval monitoring data. Fig. 13 shows the fluxes for those example datasets calculated with our models. Interestingly, the Honolua (Maui) dataset yields a very consistent envelope zero salinity intercept for the four consecutive tidal segments ($C_{gw} = 123{,}928 \pm 6413$ dpm/m^3) which suggests a temporally consistent fresh groundwater C_{gw}. While this observation supports the value of our method to determine the groundwater radon endmember and fresh SGD rates for a specific survey, the benefits of the method relative to the traditional mass balance are still difficult to quantify due to a lack of calibration data for that site. Such a calibration is difficult because of methodological limitations. Numerical groundwater flow models, which are often used for SGD calibrations (e.g., Savatier and Rocha, 2021; Schubert et al., 2019), are uncertain at the local scale (i.e., SGD exit point) because they require accurate input of the difficult-to-estimate parameters recharge and hydraulic conductivity, the latter commonly varying over several orders of magnitudes on volcanic islands (e.g., Rotzoll et al., 2007). Flow estimates from seepage meters from mapped SGD exit points may be more useful, but those exit points typically occur as clusters of karstic conduits and there may always be diffuse seepage from non-instrumented exit points that contribute to the radon-enriched buoyant plume (e.g., Glenn et al., 2013; Swarzenski et al., 2012). This causes uncertainty that should be carefully considered in the calibration of our SGD quantification method.

For the 11 tidal segments of the Dor Beach dataset (Fig. 13b), our method produced a much more variable C_{gw} (mean value: 620,167 dpm/

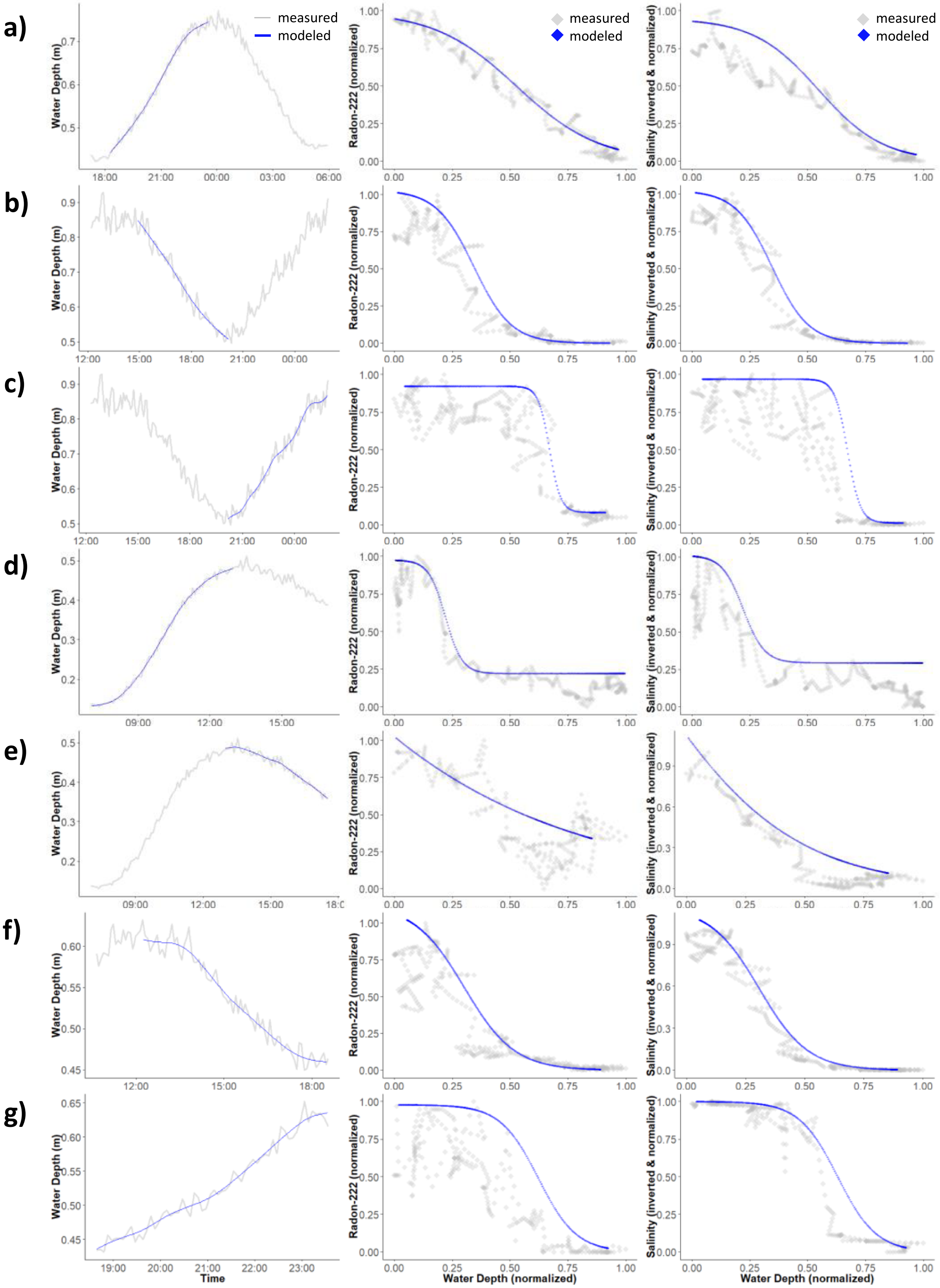


Fig. 8. Water level time series (left column) and optimized signal envelopes for radon (mid column) and salinity (right column) for different high to low tide (hl) and low to high tide (lh) segments for Varari 03/28/2022 lh **(a)**, Varari 03/21/2022 hl **(b)**, Varari 03/21/2022 lh **(c)**, Cabral 08/09/2021 lh **(d)**, Cabral 08/09/2021 hl **(e)**, Varari 08/05/2021 hl **(f)** and Varari 08/04/2021 lh **(g)**.

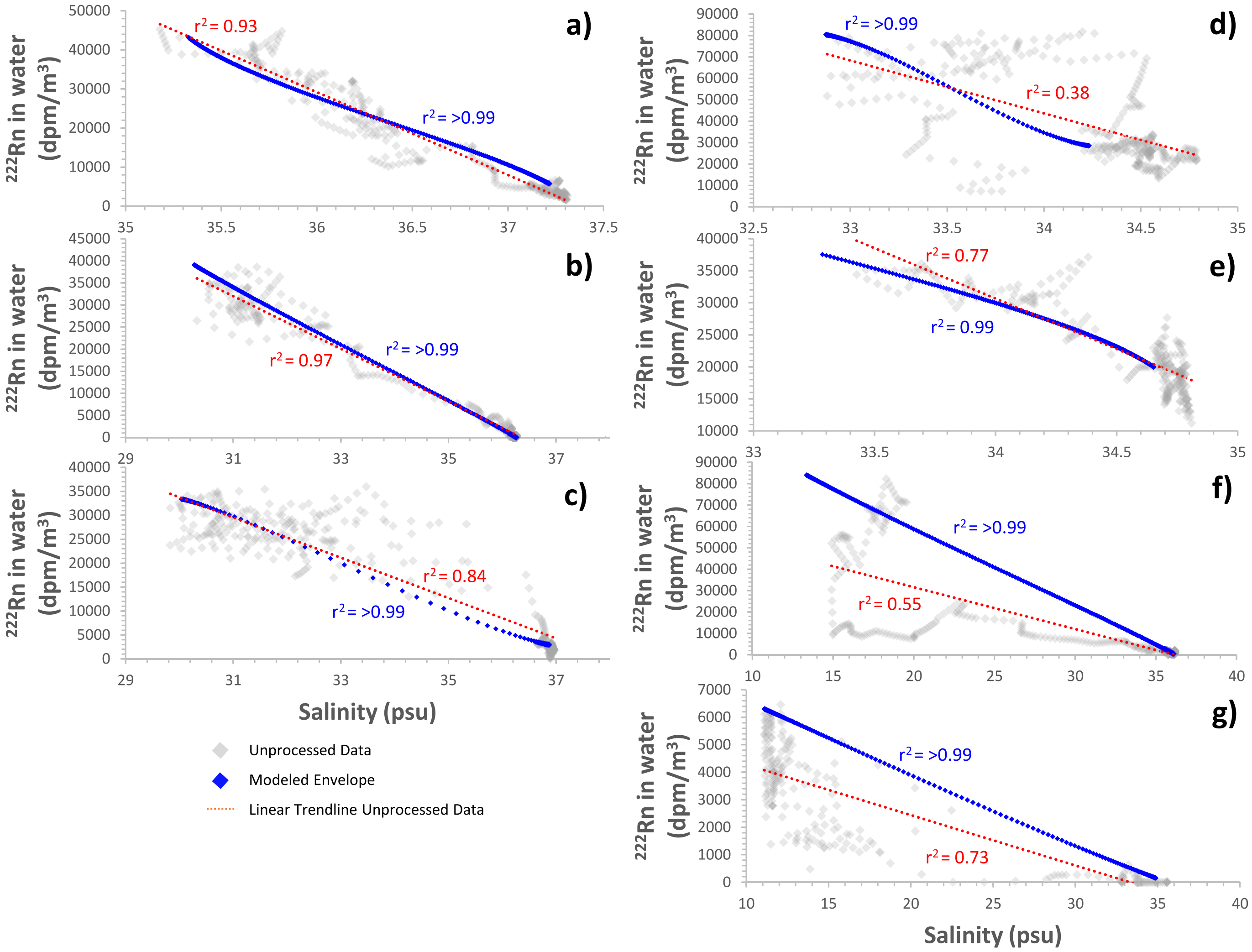


Fig. 9. Plot of correlation between measured- and denormalized envelope salinity and radon activities at Varari 03/28/2022 lh **(a)**, Varari 03/21/2022 hl **(b)**, Varari 03/21/2022 lh **(c)**, Cabral 08/09/2021 lh **(d)**, Cabral 08/09/2021 hl **(e)**, Varari 08/05/2021 hl **(f)** and Varari 08/04/2021 lh **(g)**.

Table 3
Radon and salinity data for Mo'orea groundwater samples.

Sample ID	Aquifer type	^{222}Rn (dpm/m^3)	Salinity (psu)
Coastal Spring	Carbonate	830,000	0.44
FPWD #1	Silicate	958,320	0.16
FPWD #4	Silicate	886,486	0.18
FPWD Spring #2	Silicate	395,050	0.06
FPWD #6	Silicate	1,384,614	0.13
Juice Factory Well #7	Silicate	2,169,869	0.33

m^3, one standard deviation: 361,023 dpm/m^3) that is higher than groundwater sampled in beach sand and an underlying sandstone aquifer (168,000 dpm/m^3 and 390,000 dpm/m^3, respectively; Weinstein et al., 2007a). As mentioned before, the C_{gw} variability can be related to a time variant radon and salinity signature of SGD. Importantly, the Dor Beach dataset has calibration data from seepage meter measurements available. Using the reported groundwater C_{gw} of 242,000 dpm/m^3, our mean modeled advection rate of 1.04 cm/d (one standard deviation: 1.72 cm/d) aligns much closer to the mean values from seepage meter measurements (0.2–2.4 cm/d; Weinstein et al., 2007b) than the average estimate from the traditional radon mass balance reported in that study (~8.1 cm/d; Weinstein et al., 2007b). This comparison indicates superior performance of our new method relative to the traditional radon mass balance. However, more concurrently collected SGD tracer and seepage meter datasets are necessary to further validate our method's performance.

5. Conclusions and considerations for method applications

This study demonstrates the value of nlqr, signal filtering and multi objective optimization to quantify SGD rates. Our method considers two independent tracers (radon and salinity), has a physical basis (i.e., it assumes a non-linear head loss vs discharge relationship of a sloping unconfined aquifer) and does not rely on subjective/uncertain estimates of various radon sources and sinks (e.g., F_{tide}, F_{mix} and F_{atm}). If similar envelope patterns for radon and inverted salinity vs. water level regressions can be defined, our method can quantify the fresh groundwater flux at any site according to the procedure outlined in section 2.2.

In situations where saline SGD dominates and where a monotonic regression envelope can only be defined for the radon tracer within a tidal segment, our method can still be used to estimate the saline SGD flux by (1) only modeling the radon envelope at a zero lag and high quantile (i.e., $\tau_{Rn} > 0.5$), (2) assigning C_{gw} to the maximum envelope radon value and (3) keeping the $\left(1 - \frac{S_{mon}}{S_{ref}}\right)$ term at a value of one (Eq. (1)). However, this flux should be considered a cursory estimate because of the lack of salinity data for a cross check. Appropriate radon lag, τ_{Rn} and C_{gw} cannot be objectively determined without envelope data of

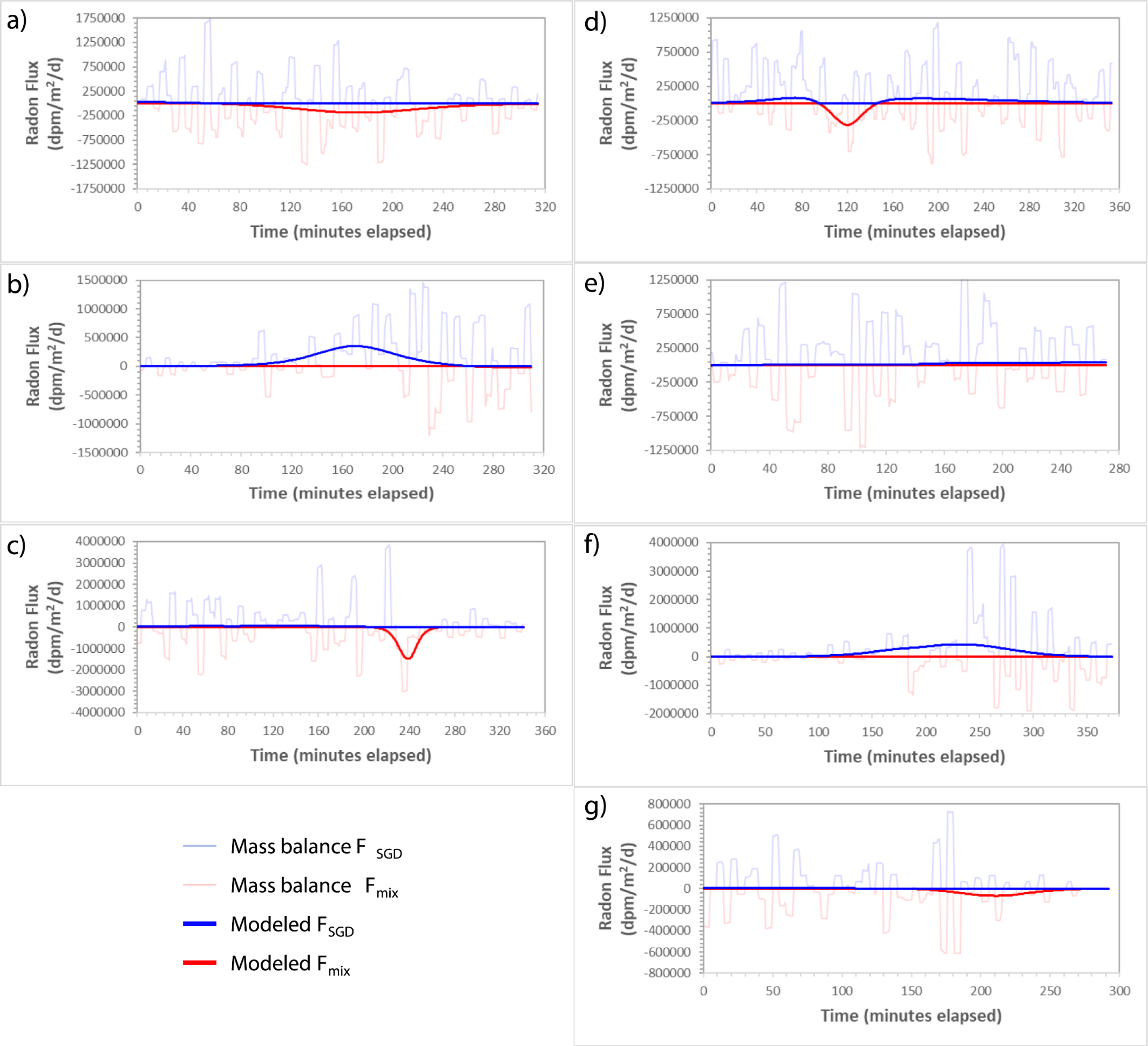


Fig. 10. Trends of F_{SGD} and F_{mix} modeled in this study and derived from the traditional mass balance applied on processed (i.e., lag corrected) monitoring data for Varari 03/28/2022 lh **(a)**, Varari 03/21/2022 hl **(b)**, Varari 03/21/2022 lh **(c)**, Cabral 08/09/2021 lh **(d)**, Cabral 08/09/2021 hl **(e)**, Varari 08/05/2021 hl **(f)** and Varari 08/04/2021 lh **(g)**.

other SGD tracers.

The situation becomes more complicated in cases where the signal shifts from a fresh to saline SGD dominance within a tidal segment. Even though our algorithm may be able to define monotonic envelopes for radon and inverted salinity in such cases, it may struggle matching them according to $F1$ (Eq. (8)). This is because fresh and saline SGD should exhibit different end-member chemistries (i.e., high radon, low salinity for fresh SGD and low-intermediate radon and high salinity for saline SGD) which will ultimately result in different envelope shapes. Therefore, in situations where the MAE of equation (8) cannot be minimized to an acceptable level (e.g., <20 %), we also recommend only modeling a radon envelope at a high (>0.5) τ_{Rn} for F_{SGD} estimation. Here, quantitative SGD assessments should only focus on the F_{SGD} radon flux (Eq. (4)) because the uniform radon and salinity chemistry assumption for SGD and zero-salinity intercept assumptions for C_{gw} in equation (1) do not apply. However, a comparison of the temporal envelope patterns of radon and inverted salinity may provide further insight on the temporal variability in radon sources from fresh and saline SGD. Still, more data-driven studies utilizing concurrent and high frequency SGD tracer- and seepage meter datasets are needed for a better differentiation between fresh SGD and saline SGD components. Given their respective differences in fluxes and geochemical characteristics, their effects on coral reef (and other) ecosystems may differ drastically.

There are some other methodological limitations that must be considered for applications to other sites. One key requirement is the availability of high-resolution monitoring data to define the envelope lines attributed to $F_{mix\ (low\ freq)}$ and F_{SGD}. If, for example, the degassing loss F_{atm} from turbulent (e.g., wave white cap) water is significant throughout the tidal segment, the algorithm may struggle to converge because radon and inverted salinity envelope shapes will differ drastically. Another important aspect to consider for future research is the dependance of the results on the type of regression function. While the

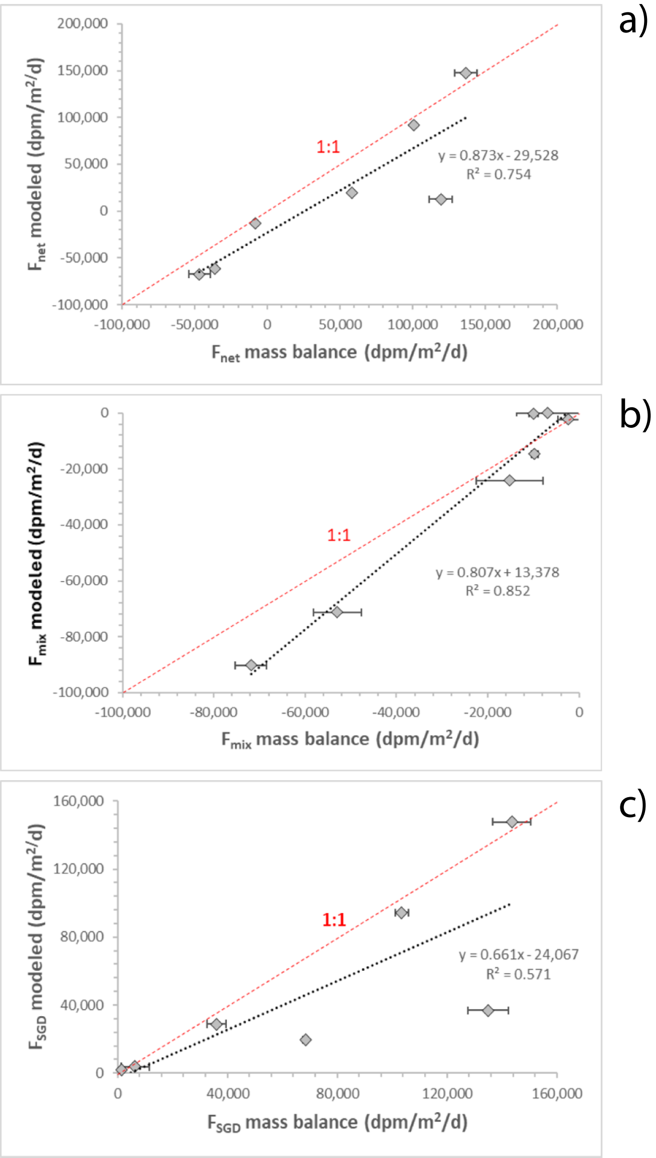


Fig. 11. Linear correlation of radon fluxes F_{net} **(a)**, F_{mix} **(b)** and F_{SGD} **(c)** modeled in this study from nlqr envelopes vs. estimated from the traditional radon mass balance. Error bars reflect mass balance uncertainty from 30- to 60-minute signal aggregations. Black dashed line is a linear regression.

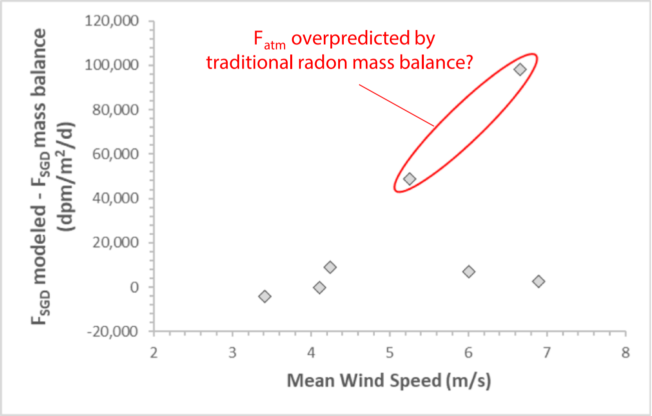


Fig. 12. Plot showing mean wind speed and difference in mean F_{SGD} rates predicted by model and estimated from traditional mass balance at a 30 to 60-minute aggregation interval.

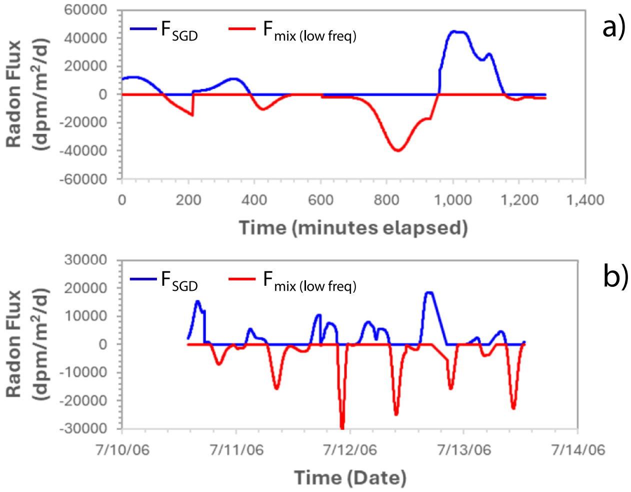


Fig. 13. Time series FSGD quantifications from nonlinear quantile regression model for (a) 4 consecutive tidal segments in Honolua (Data from Bishop et al., 2017) and (b) 11 consecutive tidal segments in Dor Beach (Weinstein et al., 2007b, 2007a).

herein applied 3pl and 4pl functions yielded better fits than other (e.g., Gompertz) sigmoidal curves, they may not be applicable to other datasets or settings. There is clearly some subjectivity in the choice of nonlinear regression function (Archontoulis and Miguez, 2015) and we recommend following our approach of considering at least two to account for some degree of the associated uncertainty. We also recommend more research on the seasonal and tidal dynamics of oceanic ^{226}Ra to provide more reliable definitions of excess radon in water activities for the regression function (see “Step 1” in Fig. 6).

Contrary to the traditional mass balance, our new method can account for the occurrences of a non-homogenous tracer inventory (i.e., stratified water column) as a mixing loss ($F_{mix\ (high\ freq)}$). This will be indicated by measured radon and salinity signal deviations from their respective envelope lines. However, if stratification persists throughout the monitoring, such as expected if monitoring occurs at some distance away from the SGD exit point (and the water extraction system captures only the shallowest part of a buoyant freshwater plume), our model will yield an overestimate of F_{SGD}. In this case, it is possible that the water extraction system captures some “legacy” radon and salinity signals from previous low tides. These signals can greatly distort the expected water level vs. radon and inverted salinity regressions per tidal segment. In these situations, chemical tracers should – if at all – only be used for a qualitative assessment of SGD.

CRediT authorship contribution statement

Benjamin Hagedorn: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Matthew W. Becker:** Supervision, Resources, Methodology, Investigation, Funding acquisition, Conceptualization. **Nyssa J. Silbiger:** Methodology, Funding acquisition, Conceptualization. **Brittney Maine:** Resources, Methodology, Investigation. **Ellen Justis:** Methodology, Investigation, Conceptualization. **Danielle M. Barnas:** Methodology, Investigation. **Maya Zeff:** Methodology, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to extend special gratitude to the staff of the University of California Berkeley Gump Research Station and La Polynésienne des Eaux water agency for logistical support in the field and groundwater well access. We are indebted to the local residents at our two field sites and the current and past indigenous stewards of Mo’orea, French Polynesia for allowing us to conduct research on their land. In particular, we thank T. Cabral, Flora, Juliette, and Bruno for graciously allowing us to work in their backyards. We also thank M. Schlüter for technical support with the MIDACO algorithm and the four anonymous reviewers, whose insightful comments greatly improved the quality of the manuscript. This research was supported by NSF grant HS-1936671 to BH and MWB and OCE-1924281 to NJS.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2024.132145.

Data availability

All data are uploaded as supplementary materials in Appendices A and B

References

Adyasari, D., Dimova, N.T., Dulai, H., Gilfedder, B.S., Cartwright, I., McKenzie, T., Fuleky, P., 2023. Radon-222 as a groundwater discharge tracer to surface waters. Earth-Sci. Rev. 238, 104321. https://doi.org/10.1016/j.earscirev.2023.104321.

Alarcon Falconi, T.M., Estrella, B., Sempértegui, F., Naumova, E.N., 2020. Effects of Data Aggregation on Time Series Analysis of Seasonal Infections. Int. J. Environ. Res. Public. Health 17, 5887. https://doi.org/10.3390/ijerph17165887.

Allen, D., Gerrans, P., Singh, A., Powell, R., 2009. Quantile regression: Its application in investment analysis. ECU Publ. 4.

Alorda-Kleinglass, A., Ruiz-Mallén, I., Diego-Feliu, M., Rodellas, V., Bruach-Menchén, J. M., Garcia-Orellana, J., 2021. The social implications of Submarine Groundwater Discharge from an Ecosystem Services perspective: A systematic review. Earth-Sci. Rev. 221, 103742. https://doi.org/10.1016/j.earscirev.2021.103742.

Archontoulis, S.V., Miguez, F.E., 2015. Nonlinear Regression Models and Applications in Agricultural Research. Agron. J. 107, 786–798. https://doi.org/10.2134/agronj2012.0506.

Arora, B., Burrus, M., Newcomer, M., Steefel, C.I., Carroll, R.W.H., Dwivedi, D., Dong, W., Williams, K.H., Hubbard, S.S., 2020. Differential C-Q Analysis: A New Approach to Inferring Lateral Transport and Hydrologic Transients Within Multiple Reaches of a Mountainous Headwater Catchment. Front, Water, p. 2.

Barillon, R., Özgümüs, A., Chambaudet, A., 2005. Direct recoil radon emanation from crystalline phases. Influence of moisture content. Geochim. Cosmochim. Acta 69, 2735–2744. https://doi.org/10.1016/j.gca.2004.11.021.

Becker, M.W., Cason, F.M., Hagedorn, B., 2023. Locating Potential Groundwater Pathways in a Fringing Reef Using Continuous Electrical Resistivity Profiling. Hydrology 10, 206. https://doi.org/10.3390/hydrology10110206.

Bergherr, E., 2018. Quantile regression: A short story on how and why. Stat. Model. 18, 1471082X1875914. doi: 10.1177/1471082X18759142.

Bishop, J.M., Glenn, C.R., Amato, D.W., Dulai, H., 2017. Effect of land use and groundwater flow path on submarine groundwater discharge nutrient flux. J. Hydrol. Reg. Stud. Water, Energy, and Food Nexus in the Asia-Pacific Region 11, 194–218. https://doi.org/10.1016/j.ejrh.2015.10.008.

Burnett, W.C., Dulaiova, H., 2003. Estimating the dynamics of groundwater input into the coastal zone via continuous radon-222 measurements. J. Environ. Radioact. 69, 21–35. https://doi.org/10.1016/S0265-931X(03)00084-5.

Cade, B.S., Noon, B.R., 2003. A gentle introduction to quantile regression for ecologists. Front. Ecol. Environ. 1, 412–420. https://doi.org/10.1890/1540-9295(2003)001[0412:AGITQR]2.0.CO;2.

Cade, B.S., 2017. Quantile Regression Applications in Ecology and the Environmental Sciences, in: Handbook of Quantile Regression. Chapman and Hall/CRC.

Cardenas, M.B., Zamora, P.B., Siringan, F.P., Lapus, M.R., Rodolfo, R.S., Jacinto, G.S., Diego-McGlone, M.L.S., Villanoy, C.L., Cabrera, O., Senal, M.I., 2010. Linking regional sources and pathways for submarine groundwater discharge at a reef by electrical resistivity tomography, 222Rn, and salinity measurements. Geophys. Res. Lett. 37. https://doi.org/10.1029/2010GL044066.

Carling, P., Jonathan, P., Su, T., 2021. Fitting limit lines (envelope curves) to spreads of geoenvironmental data. Prog. Phys. Geogr. 46. https://doi.org/10.1177/03091333211059995.

Carling, P.A., Jonathan, P., Su, T., 2022. Fitting limit lines (envelope curves) to spreads of geoenvironmental data. Prog. Phys. Geogr. Earth Environ. 46, 272–290. https://doi.org/10.1177/03091333211059995.

Cartwright, I., Gilfedder, B., Hofmann, H., 2014. Contrasts between estimates of baseflow help discern multiple sources of water contributing to rivers. Hydrol. Earth Syst. Sci. 18, 15–30. https://doi.org/10.5194/hess-18-15-2014.

Chen, X., Tang, N., Zhou, Y., 2016. Quantile regression of longitudinal data with informative observation times. J. Multivar. Anal. 144, 176–188. https://doi.org/10.1016/j.jmva.2015.11.007.

Cheng, T., Adepeju, M., 2014. Modifiable Temporal Unit Problem (MTUP) and Its Effect on Space-Time Cluster Detection. PLOS ONE 9, e100465.

Cleophas, T.J., Zwinderman, A.H., 2021. Quantile Regression in Clinical Research: Complete analysis for data at a loss of homogeneity. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-82840-0.

Colbert, S.L., Berelson, W.M., Hammond, D.E., 2008a. Radon-222 budget in Catalina Harbor, California: 2. Flow dynamics and residence time in a tidal beach. Limnol. Oceanogr. 53, 659–665.

Colbert, S.L., Hammond, D.E., Berelson, W.M., 2008b. Radon-222 Budget in Catalina Harbor, California: 1. Water Mixing Rates. Limnol. Oceanogr. 53, 651–658.

Correa, R.E., Cardenas, M.B., Rodolfo, R.S., Lapus, M.R., Davis, K.L., Giles, A.B., Fullon, J.C., Hajati, M.-C., Moosdorf, N., Sanders, C.J., Santos, I.R., 2021. Submarine Groundwater Discharge Releases CO2 to a Coral Reef. ACS EST Water 1, 1756–1764. https://doi.org/10.1021/acsestwater.1c00104.

Cyronak, T., Santos, I.R., Erler, D.V., Maher, D.T., Eyre, B.D., 2014. Drivers of pCO2 variability in two contrasting coral reef lagoons: The influence of submarine groundwater discharge. Glob. Biogeochem. Cycles 28, 398–414. https://doi.org/10.1002/2013GB004598.

deAyala, R., 2021. The Theory and Practice of Item Response Theory, 2nd ed. Guilford Press, New York, NY.

Duchesne, M.J., Gaillot, P., 2011. Did you smooth your well logs the right way for seismic interpretation? J. Geophys. Eng. 8, 514–523. https://doi.org/10.1088/1742-2132/8/4/004.

Eilers, P.H.C., 2003. A Perfect Smoother. Anal. Chem. 75, 3631–3636. https://doi.org/10.1021/ac034173t.

Fang, Y., Zheng, T., Guo, B., Zhan, H., Wang, H., Zheng, X., Walther, M., 2022. Transformation in the Stability of Tide-Induced Upper Saline Plume Driven by Transient External Forcing. Water Resour. Res. 58, e2021WR031331. doi: 10.1029/2021WR031331.

Fefferman, N.H., O’Neil, E.A., Naumova, E.N., 2005. Confidentiality and Confidence: Is Data Aggregation a Means to Achieve Both? J. Public Health Policy 26, 430–449. https://doi.org/10.1057/palgrave.jphp.3200029.

Geng, X., Heiss, J.W., Michael, H.A., Li, H., Raubenheimer, B., Boufadel, M.C., 2021. Geochemical fluxes in sandy beach aquifers: Modulation due to major physical stressors, geologic heterogeneity, and nearshore morphology. Earth-Sci. Rev. 221, 103800. https://doi.org/10.1016/j.earscirev.2021.103800.

Geraci, M., 2019. Modelling and estimation of nonlinear quantile regression with clustered data. Comput. Stat. Data Anal. 136, 30–46. https://doi.org/10.1016/j.csda.2018.12.005.

Glenn, C., Whittier, R., Dailer, M.L., Dulai, H., El-Kadi, A., Fackrell, J., Kelly, J., Waters, C., Sevadjian, J., 2013. Lahaina Groundwater Tracer Study Lahaina, Maui, Hawaiʻi, Final Report.

Grande, E., Seybold, E.C., Tatariw, C., Visser, A., Braswell, A., Arora, B., Birgand, F., Haskins, J., Zimmer, M., 2023. Seasonal and tidal variations in hydrologic inputs drive salt marsh porewater nitrate dynamics. Hydrol. Process. 37, e14951.

Hagedorn, B., 2020. Hydrograph separation through multi objective optimization: Revealing the importance of a temporally and spatially constrained baseflow solute source. J. Hydrol. 590, 125349. https://doi.org/10.1016/j.jhydrol.2020.125349.

Hagedorn, B., Becker, M.W., Silbiger, N.J., 2020. Evidence of freshened groundwater below a tropical fringing reef. Hydrogeol. J. 28, 2501–2517. https://doi.org/10.1007/s10040-020-02191-1.

Hagedorn, B., Tsuda, M., 2022. Radon and Salinity Mass Balance Constraints on Groundwater Recharge on a Semi-Arid Island (Catalina, California). Water 14, 1068. https://doi.org/10.3390/w14071068.

Hao, L., Naiman, D.Q., 2007. Quantile Regression. SAGE.

Haßler, K., Dähnke, K., Kölling, M., Sichoix, L., Nickl, A.-L., Moosdorf, N., 2019. Provenance of nutrients in submarine fresh groundwater discharge on Tahiti and Moorea, French Polynesia. Appl. Geochem. 100, 181–189. https://doi.org/10.1016/j.apgeochem.2018.11.020.

Houben, G.J., Stoeckl, L., Mariner, K.E., Choudhury, A.S., 2018. The influence of heterogeneity on coastal groundwater flow - physical and numerical modeling of fringing reefs, dykes and structured conductivity fields. Adv. Water Resour. 113, 155–166. https://doi.org/10.1016/j.advwatres.2017.11.024.

Johnson, A.G., Glenn, C.R., Burnett, W.C., Peterson, R.N., Lucey, P.G., 2008. Aerial infrared imaging reveals large nutrient-rich groundwater inputs to the ocean. Geophys. Res. Lett. 35. https://doi.org/10.1029/2008GL034574.

Justus, A.L., 2019. Multiple Facets of the Poisson Distribution Applicable to Health Physics Measurements. Health Phys. 117, 36. https://doi.org/10.1097/HP.0000000000001013.

Kim, G., Burnett, W.C., Dulaiova, H., Swarzenski, P.W., Moore, W.S., 2001. Measurement of 224Ra and 225Ra activities in natural waters using a radon-in-air monitor. Environ. Sci. Technol. 35, 4680–4683. https://doi.org/10.1021/es010804u.

Knee, K.L., Crook, E.D., Hench, J.L., Leichter, J.J., Paytan, A., 2016. Assessment of Submarine Groundwater Discharge (SGD) as a Source of Dissolved Radium and Nutrients to Moorea (French Polynesia) Coastal Waters. Estuaries Coasts 39, 1651–1668. https://doi.org/10.1007/s12237-016-0108-y.

Koenker, R., 2017. Quantile Regression: 40 Years On. Annu. Rev. Econ. 9, 155–176.

Koenker, R., Park, B.J., 1996. An interior point algorithm for nonlinear quantile regression. J. Econom. 71, 265–283. https://doi.org/10.1016/0304-4076(96)84507-6.

Kreyns, P., Geng, X., Michael, H.A., 2020. The influence of connected heterogeneity on groundwater flow and salinity distributions in coastal volcanic aquifers. J. Hydrol. 586, 124863. https://doi.org/10.1016/j.jhydrol.2020.124863.

Kuan, W.K., Xin, P., Jin, G., Robinson, C.E., Gibbes, B., Li, L., 2019. Combined Effect of Tides and Varying Inland Groundwater Input on Flow and Salinity Distribution in Unconfined Coastal Aquifers. Water Resour. Res. 55, 8864–8880. https://doi.org/10.1029/2018WR024492.

Lamontagne, S., Le Gal La Salle, C., Hancock, G.J., Webster, I.T., Simmons, C.T., Love, A. J., James-Smith, J., Smith, A.J., Kämpf, J., Fallowfield, H.J., 2008. Radium and radon radioisotopes in regional groundwater, intertidal groundwater, and seawater in the Adelaide Coastal Waters Study area: Implications for the evaluation of submarine groundwater discharge. Mar. Chem., Measurement of Radium and Actinium Isotopes in the marine environment 109, 318–336. doi: 10.1016/j.marchem.2007.08.010.

Lê Cook, B., Manning, W.G., 2013. Thinking beyond the mean: a practical guide for using quantile regression methods for health services research. Shanghai Arch. Psychiatry 25, 55–59. https://doi.org/10.3969/j.issn.1002-0829.2013.01.011.

Lloyd, C.E.M., Freer, J.E., Johnes, P.J., Collins, A.L., 2016. Using hysteresis analysis of high-resolution water quality monitoring data, including uncertainty, to infer controls on nutrient and sediment transfer in catchments. Sci. Total Environ. 543, 388–404. https://doi.org/10.1016/j.scitotenv.2015.11.028.

Luo, M., Zhang, Y., Li, H., Wang, X., Xiao, K., 2020. Submarine Groundwater Discharge in a Coastal Bay: Evidence from Radon Investigations. Water 12, 2552. https://doi.org/10.3390/w12092552.

Michael, H.A., Scott, K.C., Koneshloo, M., Yu, X., Khan, M.R., Li, K., 2016. Geologic influence on groundwater salinity drives large seawater circulation through the continental shelf. Geophys. Res. Lett. 43, 10782–10791. https://doi.org/10.1002/2016GL070863.

Moosdorf, N., Stieglitz, T., Waska, H., Dürr, H., Hartmann, J., 2015. Submarine groundwater discharge from tropical islands: a review. Grundwasser 20, 53–67. https://doi.org/10.1007/s00767-014-0275-3.

Muggeo, V., 2016. Segmented Mixed Models with Random Changepoints in r. https://doi.org/10.13140/RG.2.1.4180.8402.

Null, K.A., Knee, K.L., Crook, E.D., de Sieyes, N.R., Rebolledo-Vieyra, M., Hernández-Terrones, L., Paytan, A., 2014. Composition and fluxes of submarine groundwater along the Caribbean coast of the Yucatan Peninsula. Cont. Shelf Res. 77, 38–50. https://doi.org/10.1016/j.csr.2014.01.011.

Oehler, T., Bakti, H., Lubis, R.F., Purwoarminta, A., Delinom, R., Moosdorf, N., 2019. Nutrient dynamics in submarine groundwater discharge through a coral reef (western Lombok, Indonesia). Limnol. Oceanogr. 64, 2646–2661. https://doi.org/10.1002/lno.11240.

Pain, A.J., Martin, J.B., Young, C.R., Valle-Levinson, A., Mariño-Tapia, I., 2020. Carbon and phosphorus processing in a carbonate karst aquifer and delivery to the coastal ocean. Geochim. Cosmochim. Acta 269, 484–495. https://doi.org/10.1016/j.gca.2019.10.040.

Petermann, E., Schubert, M., 2015. Quantification of the response delay of mobile radon-in-air detectors applied for detecting short-term fluctuations of radon-in-water concentrations. Eur. Phys. J. Spec. Top. 224, 697–707. https://doi.org/10.1140/epjst/e2015-02400-5.

Peterson, R.N., Burnett, W.C., Glenn, C.R., Johnson, A.G., 2009. Quantification of point-source groundwater discharges to the ocean from the shoreline of the Big Island. Hawaii. Limnol. Oceanogr. 54, 890–904. https://doi.org/10.4319/lo.2009.54.3.0890.

Richardson, C.M., Dulai, H., Popp, B.N., Ruttenberg, K., Fackrell, J.K., 2017. Submarine groundwater discharge drives biogeochemistry in two Hawaiian reefs. Limnol. Oceanogr. 62, S348–S363. https://doi.org/10.1002/lno.10654.

Robinson, C.E., Xin, P., Santos, I.R., Charette, M.A., Li, L., Barry, D.A., 2018. Groundwater dynamics in subterranean estuaries of coastal unconfined aquifers: Controls on submarine groundwater discharge and chemical inputs to the ocean. Adv. Water Resour. 115, 315–331. https://doi.org/10.1016/j.advwatres.2017.10.041.

Rostami-Tabar, B., Mircetic, D., 2023. Exploring the association between time series features and forecasting by temporal aggregation using machine learning. Neurocomputing 548, 126376. https://doi.org/10.1016/j.neucom.2023.126376.

Rotzoll, K., El-Kadi, A.I., Gingerich, S.B., 2007. Estimating Hydraulic Properties of Volcanic Aquifers Using Constant-Rate and Variable-Rate Aquifer Tests1. JAWRA J. Am. Water Resour. Assoc. 43, 334–345. https://doi.org/10.1111/j.1752-1688.2007.00026.x.

Roy, I.G., 2020. An optimal Savitzky-Golay derivative filter with geophysical applications: an example of self-potential data. Geophys. Prospect. 68, 1041–1056. https://doi.org/10.1111/1365-2478.12892.

Rupp, D.E., Selker, J.S., 2006. On the use of the Boussinesq equation for interpreting recession hydrographs from sloping aquifers. Water Resour. Res. 42. https://doi.org/10.1029/2006WR005080.

Santos, I.R., Erler, D., Tait, D., Eyre, B.D., 2010. Breathing of a coral cay: Tracing tidally driven seawater recirculation in permeable coral reef sediments. J. Geophys. Res. Oceans 115. https://doi.org/10.1029/2010JC006510.

Santos, I.R., Maher, D.T., Eyre, B.D., 2012. Coupling Automated Radon and Carbon Dioxide Measurements in Coastal Waters. Environ. Sci. Technol. 46, 7685–7691. https://doi.org/10.1021/es301961b.

Savatier, M., Rocha, C., 2021. Rethinking tracer-based (Ra, Rn, salinity) approaches to estimate point-source submarine groundwater discharge (SGD) into coastal systems. J. Hydrol. 598, 126247. https://doi.org/10.1016/j.jhydrol.2021.126247.

Schlueter, M., 2014. MIDACO software performance on interplanetary trajectory benchmarks. Adv. Space Res. 54, 744–754. https://doi.org/10.1016/j.asr.2014.05.002.

Schlueter, M., Erb, S.O., Gerdts, M., Kemble, S., Rückmann, J.-J., 2013. MIDACO on MINLP space applications. Adv. Space Res. 51, 1116–1131. https://doi.org/10.1016/j.asr.2012.11.006.

Schlüter, M., Egea, J.A., Banga, J.R., 2009. Extended ant colony optimization for non-convex mixed integer nonlinear programming. Comput. Oper. Res. 36, 2217–2229. https://doi.org/10.1016/j.cor.2008.08.015.

Schmid, M., Rath, D., Diebold, U., 2022. Why and How Savitzky-Golay Filters Should Be Replaced. ACS Meas. Sci. Au 2, 185–196. https://doi.org/10.1021/acsmeasuresciau.1c00054.

Schubert, M., Paschke, A., Lieberman, E., Burnett, W.C., 2012. Air-Water Partitioning of 222Rn and its Dependence on Water Temperature and Salinity. Environ. Sci. Technol. 46, 3905–3911. https://doi.org/10.1021/es204680n.

Schubert, M., Petermann, E., Stollberg, R., Gebel, M., Scholten, J., Knöller, K., Lorz, C., Glück, F., Riemann, K., Weiß, H., 2019. Improved Approach for the Investigation of Submarine Groundwater Discharge by Means of Radon Mapping and Radon Mass Balancing. Water 11, 749. https://doi.org/10.3390/w11040749.

Schubert, M., Scholten, J., Kreuzburg, M., Petermann, E., de Paiva, M.L., Köhler, D., Liebetrau, V., Rapaglia, J., Schlüter, M., 2022. Radon (222Rn) as tracer for submarine groundwater discharge investigation—limitations of the approach at shallow wind-exposed coastal settings. Environ. Monit. Assess. 194, 798. https://doi.org/10.1007/s10661-022-10462-5.

Silbiger, N.J., Donahue, M.J., Lubarsky, K., 2020. Submarine groundwater discharge alters coral reef ecosystem metabolism. Proc. r. Soc. B Biol. Sci. 287, 20202743. https://doi.org/10.1098/rspb.2020.2743.

Stieglitz, T.C., Cook, P.G., Burnett, W.C., 2010. Inferring coastal processes from regional-scale mapping of 222Radon and salinity: examples from the Great Barrier Reef. Australia. J. Environ. Radioact. 101, 544–552. https://doi.org/10.1016/j.jenvrad.2009.11.012.

Su, N., Burnett, W.C., MacIntyre, H.L., Liefer, J.D., Peterson, R.N., Viso, R., 2014. Natural Radon and Radium Isotopes for Assessing Groundwater Discharge into Little Lagoon, AL: Implications for Harmful Algal Blooms. Estuaries Coasts 37, 893–910.

Swarzenski, P.W., Storlazzi, C.D., Presto, M.K., Gibbs, A.E., Smith, C.G., Dimova, N.T., Dailer, M.L., Logan, J.B., 2012. Nearshore morphology, benthic structure, hydrodynamics, and coastal groundwater discharge near Kahekili Beach Park, Maui, Hawaii (No. 2012–1166), Open-File Report. U.S. Geological Survey. doi: 10.3133/ofr20121166.

Tait, D.R., Santos, I.R., Erler, D.V., Befus, K.M., Cardenas, M.B., Eyre, B.D., 2013. Estimating submarine groundwater discharge in a South Pacific coral reef lagoon using different radioisotope and geophysical approaches. Mar. Chem. Radium and Radon Tracers in Aquatic Systems 156, 49–60. https://doi.org/10.1016/j.marchem.2013.03.004.

Taniguchi, M., Dulai, H., Burnett, K.M., Santos, I.R., Sugimoto, R., Stieglitz, T., Kim, G., Moosdorf, N., Burnett, W.C., 2019. Submarine Groundwater Discharge: Updates on Its Measurement Techniques, Geophysical Drivers, Magnitudes, and Effects. Front. Environ, Sci, p. 7.

Wang, X., Li, H., Yang, J., Zheng, C., Zhang, Y., An, A., Zhang, M., Xiao, K., 2017. Nutrient inputs through submarine groundwater discharge in an embayment: A radon investigation in Daya Bay, China. J. Hydrol. Investigation of Coastal Aquifers 551, 784–792. https://doi.org/10.1016/j.jhydrol.2017.02.036.

Weinert, H.L., 2007. Efficient computation for Whittaker-Henderson smoothing. Comput. Stat. Data Anal. 52, 959–974. https://doi.org/10.1016/j.csda.2006.11.038.

Weinstein, Y., Burnett, W.C., Swarzenski, P.W., Shalem, Y., Yechieli, Y., Herut, B., 2007a. Role of aquifer heterogeneity in fresh groundwater discharge and seawater recycling: An example from the Carmel coast. Israel. J. Geophys. Res. Oceans 112. https://doi.org/10.1029/2007JC004112.

Weinstein, Y., Shalem, Y., Burnett, W.C., Swarzenski, P., Herut, B., 2007b. Temporal variability of submarine groundwater discharge: Assessments via radon and seep meters, the southern carmel coast. Israel. IAHS-AISH Publ. 125–133.

Wolfe, W.W., Murgulet, D., Gyawali, B., Sterba-Boatwright, B., 2023. Modeling time series radon inventory and constraints on the submarine groundwater discharge mass balance of a well-mixed, highly dynamic estuary. J. Hydrol. 625, 130065. https://doi.org/10.1016/j.jhydrol.2023.130065.

Wymore, A.S., Leon, M.C., Shanley, J.B., McDowell, W.H., 2019. Hysteretic Response of Solutes and Turbidity at the Event Scale Across Forested Tropical Montane Watersheds. Front. Earth Sci. 7.

Xin, P., Wang, S.S.J., Lu, C., Robinson, C., Li, L., 2015. Nonlinear interactions of waves and tides in a subterranean estuary. Geophys. Res. Lett. 42, 2277–2284. https://doi.org/10.1002/2015GL063643.

Yu, X., Xin, P., Lu, C., Robinson, C., Li, L., Barry, D.A., 2017. Effects of episodic rainfall on a subterranean estuary. Water Resour. Res. 53, 5774–5787. https://doi.org/10.1002/2017WR020809.

Zhao, S., Li, M., Burnett, W.C., Cheng, K., Li, C., Guo, J., Yu, S., Liu, W., Yang, T., Dimova, N.T., Chen, G., Yu, Z., Xu, B., 2022. In-situ radon-in-water detection for high resolution submarine groundwater discharge assessment. Front. Mar, Sci, p. 9.

Zobaa, A., 2019. Mixed-Integer Distributed Ant Colony Multi-Objective Optimization of Single-Tuned Passive Harmonic Filter Parameters. IEEE Access PP, 1–1. doi: 10.1109/ACCESS.2019.2903910.