Thank you to the reviewers for taking the time to review the paper and for your supportive comments, feedback and suggestions. These have provided useful improvements to the paper, and I address each one below.

Reviewer 1:

Comment (1)

L33–36: The authors jump directly to channel rip currents here, whereas at this stage I think the text should remain more general. Rip currents can arise from a range of mechanisms and do not necessarily require alongshore gradients in wave setup (e.g. some headland rip currents). I would simply state that rip currents are generated by the action of breaking waves and move the discussion of this specific mechanism to the third paragraph.

I agree with your suggestion. The first paragraph (lines 33-35) now simply states:

> 'Rip currents are wave-driven flows that originate in the surfzone of wave-exposed beaches and can carry water users from the shallows out into deeper water where panic, exhaustion, and in some cases drowning, can occur (Brander et al., 2011; Cornell et al., 2023).'

And lines 64-67 now have the information removed from the first paragraph:

> 'Wave breaking over the sandbars creates a reduction in momentum flux, resulting in wave induced set up at the shore (Longuet-Higgins and Stewart, 1964). Gradients in the set up drive flows alongshore which return back out to sea in concentrated and often fast-moving, O(0.5-1 m s-1), offshore flows (Brander, 1999; MacMahan et al., 2006).'

Comment (2)

L49: I would not cite this equation here.

Thanks for this suggestion, I have now removed that citation from line 58-59. It now reads:

> '…defined using rip incident totals normalized by surf zone population in the UK over the 5-year period they studied.'

Comment (3)

L83 (and elsewhere): Castelle et al. (2024, the discussion paper) should be updated to 2025 (the published paper).

Thanks for pointing this out. I have now updated all the 2024 references to the published 2025 version.

Comment (4)

L99 (and elsewhere): "Sect. 0" appears in several places. Please fix the section cross-referencing.

Thanks for alerting us of this, I have now fixed these references by manually typing in the section numbers.

Comment (5)

L113–115: Since you explain the dominant wave regime mechanism for the southwest-facing coast, it would be useful to do the same for this part of the coastline. I assume it is primarily exposed to trade-wind waves, with occasional tropical cyclone swells explaining the reported 8 m extremes.

I have now updated the text to include the following on lines 133-136:

> 'The north-east facing coastline is the most sheltered overall, receiving locally-generated windseas and less frequent swells from tropical cyclones in the southwest Pacific and subtropical Pacific trade-wind belt (Laing, 2000), resulting in mean and maximum significant wave heights at the coast of 1 m and 8 m, respectively, and mean zero-crossing period of 6 s (Albuquerque et al., 2021; Gorman et al., 2003).'

Comment (6)

Figure 1: I would use different symbols and/or colours to distinguish dissipative, intermediate, and reflective beaches. This would immediately show the distribution of beach states, information that is only touched upon much later in the manuscript. For example, it would be interesting to see the relative proportion of each beach state, whether reflective and dissipative beaches are predominantly located on east- and west-facing coasts, respectively, and how well-known sites such as Muriwai are classified. Section 2.1 could then be expanded slightly to describe these patterns.

We have now added three different symbols to figure 1 to depict the three main beach types, have added key place names, and have added the beach classification of all 92 studied beaches to the table in Appendix A. I have added some text to Section 2.1 to summarise the spatial distribution of these beach types on lines 142-147:

> 'Fig. 1 demonstrates the spatial distribution of the three main beach morphology types (dissipative, intermediate, reflective) at the beaches studied in this paper. The exposed west and east facing coastlines are dominated by intermediate 'bar-rip' beach morphology, synonymous with rip currents (for example, Muriwai Beach), while wave-sheltered, tide-dominated locations exhibit dissipative morphology (for example, Tahunanui Beach), and semi-sheltered beaches with moderate tide ranges and coarser or denser sediments exhibit reflective morphology (for example, Mairangi Bay).'

Comment (7)

L50: Wow, this is the kind of dataset one can only dream about. However, I may have missed something, but this sentence suggests hourly estimates, whereas later (e.g. L405, "4.3 assessments per day on average") it seems different, i.e. with estimates not necessarily collected throughout the supervising hours. Please clarify.

Yes, it has been a great dataset to work with! On L152 we refer to the routine water user headcounts made by lifeguards at nearly all SLSNZ beaches, whereas on L432 we refer to rip current hazard perceptions made at only 12 beaches for this present study. I hope these are distinct enough in their descriptions, but happy to make this clearer if the reviewer thinks that is necessary.

Comment (8)

L177: "Only forcing combinations associated with at least 20 hours of lifeguard observations were used to estimate Prip." Given that Prip is computed using 2-hour bins, I am not sure I understand this sentence. Could you clarify exactly what is meant here?

Your comment has made me realise that this sentence is now redundant as it relates to a part of the analysis that is no longer included. It was initially included when we were planning to plot discretised bi-variate combinations of wave/tide conditions against Prip, which we didn't end up including in this paper. Instead, we assess average Prip at each Rip Index level and conditions associated with all events at Prip > 0.2. As such, L204-208 now reads as follows:

> ' P_{rip} was quantified for each beach by discretising the lifeguard data into two-hour time bins and dividing the number of incidents that occur in each time period by the instantaneous number of water users counted by the lifeguards. A subset of 'high-hazard' events was also analysed, defined as conditions where P_{rip} exceeds the 90^{th}%ile; equivalent to P_{rip} = 0.2.'

Comment (9)

L178–179: "1 rip-related incident per 5 people in the water" as a "high hazard" event seems extremely high to me. It made me wonder how beach lifeguarding and supervised bathing operate in New Zealand. For example, are rescues primarily associated with swim-between-the-flags areas? Do incidents occur both inside and outside the flagged bathing zones, and if so, in what approximate proportions? Is the red flag (no swimming) sometimes displayed, and if so, are those days excluded from the analysis to avoid biasing the model? Were some (all?) of the Prip>0.2 related to mass rescue events ? Including this type of information in Section 2 would help readers interpret the results.

I've now added the following text on L207-214:

> 'A subset of 'high-hazard' events was also analysed, defined as conditions where P_{rip} exceeds the 90^{th}%ile; equivalent to P_{rip} = 0.2. Only 448 of these events exist in the dataset and they occurred either when a single person got into a rip incident with only 5 people counted in the water or when a mass rescue event occurred with, for example, 10 simultaneous incidents with 50 people counted in the water. The data shows that the median n_{exp} during these high-hazard events was 5 people with an interquartile range of 1-13 people and the median n_{inc} during high-hazard events was 3 incidents with an interquartile range of 1-8 incidents (at one location in a single two-hour time period). Conversely, during 'standard' rip incidents (P_{rip} < 0.2), n_{exp} was 53 people with an interquartile range of 27-105 people and the median n_{inc} was 1 incident with an interquartile range of 1-2 incidents.'

Regarding your question about bathing area, incidents can occur inside and outside the flagged area, but the data does not differentiate it. Therefore, the proportion is unknown. The red flag is a rare event, and SLSNZ don't seem to capture red flag events in the data either.

Comment (10)

L254: Lifeguard perception on a 1–5 scale. I am wondering whether this scale reflects a hazard assessment normalised to each individual beach. For example, at a very sheltered beach where 0.5 m waves represent the most energetic conditions, would that be rated 5/5, whereas the same surf conditions at Muriwai would be rated 1/5? Or is this intended as a more general, *transferable assessment of hazard? From the later discussion, it seems to be the latter, but it would be useful at this stage of the paper to explain how lifeguards were instructed to use this scale.*

The 1-5 scale was intended to be transferable between different beaches. I have tried to clarify this on lines 294-298, and have now included acknowledgment of the limitations of including wave descriptions in the reference information provided to lifeguards:

> 'Lifeguards were asked to subjectively rate the rip current hazard between 1 and 5 and were provided with the rip warning levels and forcing descriptions in **Error! Reference source not found.** as a reference. This was necessary to ensure a transferable scale between beaches with different levels of wave exposure, such as a sheltered beach where 0.5 m waves represent the most energetic conditions and a more energetic beach where waves frequently exceed 2 m. However, it is acknowledged that this information may have influenced (consciously or unconsciously) the perceived rip hazard level.'

Comment (11)

Sections 3.2 and 3.3 : Tables and figure are great and insightful by provide average information and metrics, but I would love to see example of time series, ideally for three representatives beaches of the three different beach states, including long time series and maybe zooms on specific events. I understand that this may sound a bit useless, but as a reader I would really like to see how the rescue data is distributed and how the algorithm behaves to have a better understanding

I agree that this would make a good addition and illustrate the behaviour of the algorithm and distribution of incidents. I will plot an example time series with incidents (stars), RI level, Hs and Tfac, for an example location.

Comment (12)

Thankyou! I was pleased with how this figure turned out. I have updated the colour of RI1 in Fig. 6 and Fig. 9, as green does not work well for colour blindness.

KS to update colours in Fig 9 and Fig 6

Reviewer 2:

Comment (1)

I thought the Introduction was good, but needed some additional information. I felt that the overall significance of rip currents as a hazard in terms of fatalities (and in comparison with other beach/coastal hazards) was understated. I also thought more discussion of the nationally broadcast rip hazard system in the US needed more attention as this is the longest running rip current forecast of its kind. There also needs to be more explanation/justification about why NZ is the chosen location (aside from the excellent dataset). How much of a problem are rip currents in NZ – what is the context? The aims/objectives should be followed by a clear statement of the potential benefits/outcomes of this study…expand more on 'management and mitigation of the rip current hazard in NZ'.

Thanks for these suggestions. I have added the following text to the opening paragraph of the introduction on lines 35-44 to address your first point above about significance of rip current hazard, and your third point about the importance of rip hazards in NZ:

> 'International rip current incident and fatality reporting is inconsistent, but available statistics suggest that rip currents cause more than 50% of lifeguard rescues on surf beaches (Brewster et al., 2019; Brighton et al., 2013; Woodward et al., 2013) and fatalities numbering approximately 100/year in the United States (Brewster et al., 2019), 21/year in Australia (Kelly et al., 2025), and >10/year (including fatal and severe non-fatal drowning) in southwest France (Castelle et al., 2020). As in other countries, rip currents account for more than half of all lifeguard rescues in New Zealand (Sect. 3.1) and police and coroner's reports attributed rip currents to 17% (29 out of 170) of drowning deaths at surf beaches in New Zealand between January 2014 and June 2024 (Water Safety New Zealand, 2024), although this is expected to be an underestimate due to the challenges of attribution after the drowning has occurred. Rip Currents are therefore one of the primary life risks faced by water-users at beaches in New Zealand.'

I have added the following text to lines 85-87 of the introduction to emphasise the importance of the rip hazard system in the US:

> 'This was further validated (Churma et al., 2017; Gibbs et al., 2015; Moulton et al., 2017) before becoming the first publicly-broadcast, national-scale rip current forecast system in the world. Another long-running (2014-present) national-scale rip forecast was developed in the UK, based on...'

I have added the following to lines 102-104 to expand more on 'management and mitigation of the rip current hazard in NZ':

> 'It is hoped that such a forecast will inform lifeguard management decisions such as rostering ahead of high-hazard times and inform water users of the risk posed by rips and inform decisions about where to bathe on a given day. The objectives of the research are to...'

Comment (2)

The dataset, methods and results are done very well, but I did find the Rip Index Levels to be a bit limited in terms of actual rip characteristics (type, number, size, flow strength) although perhaps this is a recommendation for future studies to consider. The discussion provided a solid comparison with the literature, but I think a more nuanced section on what this forecasting tool might look like to the public would have been useful? Is informing lifeguards and lifeguard managers of times of peak rip current hazard something that they wouldn't already know through experience? Some recommendations for future research into evaluating the efficacy of a forecast system would be useful.

Thank you for these very valid suggestions. Regarding including more information on rip types, number, size, and flow strengths with the rip indexes, these things were not explicitly measured by this study and would require subjectivity to assign to each RI level. Arguably rip type, number, and size are poorly constrained within each RI level and considerable overlap would be expected. Flow strength is perhaps the most well understood in terms of its relation to forcing conditions and this could feasibly be estimated for each RI level from existing literature, but this could lead people to assume that we've quantified it at each RI level in this study, which we have not. I agree that suggesting this for future research is a valid suggestion and I've added the following to lines 472-475:

> '...While this provides good quantification of life risk and allows physical hazard to be inferred from the normalised incident numbers, future research could seek to further improve the Rip Index levels by associating each level with contextual data such as indicative flow speeds from in-situ measurements, or rip current types, number, and size from in-situ or remote (i.e. aerial or satellite) observations.'

Regarding what the forecasting tool might look like to the public and/or lifeguards we've provided some information in the discussion already on how we expect the warnings to be used and have now added some information on how the warnings might be presented and how their efficacy might be evaluated in future on lines 560-566:

> 'Future research should evaluate how the Rip Index is best communicated to different end users. While experienced lifeguards at familiar beaches may already recognise periods of elevated rip hazard, the forecast could provide additional decision support for novice lifeguards or personnel deployed to unfamiliar locations. For the public, a simple beach-specific 1–5 hazard scale paired with consistent behavioural guidance may be more appropriate than treating each elevated value as a discrete warning, particularly at exposed beaches where high hazard levels may occur frequently. Evaluation of a public-facing system should therefore extend beyond forecast skill to include warning comprehension, perceived risk, trust, behavioural response, and, where feasible, changes in safety outcomes.'

Comment (3)

L9 – a small point, but instead of 'wave-exposed coasts' I would say 'wave-exposed beaches'

Now changed to say 'wave-exposed beaches' on line 9:

> 'Rip currents are dangerous flows in the surfzone of wave-exposed beaches and can take bathers from the...'

Comment (4)

L18 – would be useful to provide a value for the average breaker height and average tide level for context

The average values are specific to each beach and provide a means to normalise the data between beaches. Therefore, it's not possible to succinctly include these data in the abstract. I've now clarified this on lines 18-19:

> '...and occurred disproportionately during wave conditions at or above average breaker height with tide level at or below average low tide for a given beach'

Comment (5)

L39 – I think it's worth having a simple statement describing different types of rip circulation which lead to both surf zone retention and exits…as this terminology may not be known by all readers

I have now added a description of the three main circulation types and associated retention and hazard levels on lines 46-50:

> ‘Surfzone retention is largely driven by the circulation pattern occurring, which can vary between “alongshore” flow at the shore with high retention and a low level of bathing hazard; “rotational” surfzone flows with circulating offshore and onshore currents that result in an intermediate level of retention and bathing hazard; and “exiting” surfzone currents where rips flow directly or obliquely out of the surfzone representing the lowest retention and highest hazard to bathers (Scott et al., 2014).’

Comment (6)

L64 – forewarning members of the public of what? The presence or severity of rips? Would be worth making it clear what the purpose of rip forecasts are as this is the crux of the paper

I have now added some words to clarify this on L77-78:

> ‘with the aim of forewarning members of the public that a bathing hazard exists at a given point in time or informing lifeguard rostering and operations.’

Comment (7)

L79 – South Korea

Now corrected on line 94:

> ‘…a system for four beaches in South Korea developed from…’

Comment (8)

L96 – a predictive algorithm….of what? Best to state I think.

Now corrected on line 102:

> ‘…simple predictive rip current hazard algorithm that can be operationalised within a national-scale forecast system.’

Comment (9)

L97 – so the met-ocean forcing parameters are wave, tide and wind conditions? Might want to clarify…I find this opening paragraph a bit hard to follow to be honest…too much detail too soon. Can it be simplified to just a few lines describing how the Methods section is structured/delivered?

On reflection, I agree the methods section needs some opening sentences. I’ve now re-worded lines 111-114:

> ‘This study uses a large number of lifeguard-recorded, rip-related bathing incidents and water user head counts (Sect. 2.2) to estimate rip current hazard (Sect. 2.3) and associate this with environmental forcing conditions occurring at different times (Sect 2.4). The distributions of wave, tide, and wind (herein termed met-ocean) forcing conditions occurring at incident and non-incident times are used to develop a predictive algorithm for rip current hazard.’

I’ve also added some Section numbers on lines 119-122 of that paragraph to continue to describe the structure of the methods section:

> ‘To account for water-user exposure, the algorithm was validated for its ability to differentiate times of low and high rip hazard (Sect. 3.3.1), parameterised as the

likelihood of an individual being involved in a rip incident. The predictions of rip hazard were then compared to lifeguard perceptions of rip hazard at different times and locations around New Zealand (Sect. 3.3.3).'

Comment (10)

L109 – please refer to periods as 'wave periods'

I have now checked and corrected this throughout. The word 'period' has been adjusted to 'wave period' on lines 128, 132, 135, 220, 224, 225, 330, 492, 495.

Comment (11)

Figure 1 – it would be even better if the various wave climates and tide regimes could be illustrated on this Figure (colour shading?)

I have now added wave roses and spring tide range information to figure 1, to demonstrate wave climate and tide at 5 example beaches covering each main coastline.

Comment (12)

L145 – again…met-ocean data…what do you mean by this?

As this is commonly used in offshore and coastal engineering I hadn't thought it was a jargon term, but I'm more than happy to clarify to make it more widely understandable. According to Wikipedia it is a 'syllabic abbreviation of meteorology and (physical) oceanography'. Rather than state this in the paper, I've put an indirect clarification on line 113:

> 'The distributions of wave, tide, and wind (herein termed met-ocean) forcing conditions occurring at incident and non-incident times are…'

Comment (13)

L148 – I think it should be mentioned that the head count estimates are subjective an later in limitations this subjectively should be described (e.g. different lifeguards making estimates etc.)

I have now clarified this on lines 181-182 and signposted the discussion section for lifeguard data limitations:

> 'Although it is accepted that head count data are subjective and not always accurate (Sect. 4)…'

I have also added a sentence to the paragraph on lifeguard data limitations in the discussion section (Sect. 4) on lines 520-522:

> 'When a few tens of water users are present it can be assumed that these are actual head counts of each individual in the water, but with larger numbers present these head counts will involve estimation and will therefore vary between different lifeguards, introducing subjectivity to the data.'

Comment (14)

L160-165 – what is meant by modest and highly hazardous rip current flows?

I have now clarified this on L190-194:

> ‘when only modest rip flows (i.e. with high surfzone retention and low velocity) are present… with highly hazardous rip current flows (i.e. with low surfzone retention and high velocity)…’

Comment (15)

L166 – italicise P

This is now fixed throughout by replacing Prip with P_{rip}.

Comment (16)

L185 – make sure to format variables correctly throught, e.g. H should be in italics, s should be subscript (not italic)

Thanks for flagging this. I’ve now checked and corrected accordingly throughout (too many lines to list here).

Comment (17)

L219- More detail on the satellite images is needed – what satellite images? Did the 10-20 images match the incident dates? This section is a bit fuzzy.

I’ve now added the following text to give more details on the satellite/aerial imagery used on lines 261-265:

> ‘It should be noted that these images were not intentionally chosen to coincide with times of rip current incidents. High-resolution (meter scale) imagery available in Google Earth Pro supplied by multiple providers was used for this analysis and included a mixture of commercial high-resolution satellite imagery (Digital Globe/Maxar World View, Geo Eye and Quick Bird missions), Landsat-7/8 and Sentinel-2 satellite imagery, and orthorectified aerial photography (Airbus SPOT/Pléiades, LINZ).’

Comment (18)

L245 – comma before ‘but’ please….numerous cases of this throughout

I have now added a comma where a subject and verb are present on either side of the word ‘but’; i.e. where the two parts could standalone as complete sentences. Where this is not the case it’s not appropriate to use a comma so I’ve left without a comma before the word ‘but’.

Comment (19)

- Figure 3 is not necessary

The authors disagree that this figure is not necessary and have decided to keep this in the manuscript. The figure is a relatively unique example of lifeguards collecting visual subjective hazard perception and head count data, which are central to the methods used.

Comment (20)

L256 – so lifeguards estimated rip hazard visually based on Table 1? I find it problematic that Table 1 does not actually mention characteristics about rip currents such as their presence (multiple present? Type?), width/size, and an estimate of the flow strength (only RI5 mentions strong currents)

We originally considered including information that described rip current characteristics at each Rip Index level in the table provided to lifeguards to help with their assessment. However, information such as type, number, and width/size are poorly constrained at each Rip Index level as there is a wide range of rip current behaviours that can occur at each hazard level. In the end we decided that any information we added about specific rip current characteristics at each Rip Index could be fallible and could impact the subjective perception of the lifeguards. Further, such information on rip characteristics or morphologic descriptors was not provided to lifeguards as a rating criteria in previous similar studies by Dusek and Seim (2013) or Castelle et al., (2025), who simply required their lifeguards to report their judgment of danger, intensity, or hazard level. Therefore our method was in line with that used previous studies.

I've recognised this now in some extra words added to lines 298-302:

> 'In line with previous similar studies (Castelle et al., 2025; Dusek and Seim, 2013b), lifeguards were not provided with any information on expected rip current type, behaviour, or morphologic descriptors at each Rip Index level and were simply asked to subjectively rate the rip hazard. These ratings could be provided at any time during their patrols and, during validation, the model prediction used corresponded to the closest available forecast to the time the assessment was initiated.'

Comment (21)

L274-277- I would mention the rip drownings in the Intro when NZ is established as the location of focus

I have now removed these lines from this section and have moved them to the introduction, first paragraph, as suggested. See our reply to your first comment for more details.

Comment (22)

- Don't think Figure 4 is necessary. Simple data description is sufficient

On reflection, this figure was already mostly described in the text and I agree that the figure is not necessary. I have therefore removed the figure and expanded the description in the text on lines 316-321 to capture the information that was in the figure:

> 'The majority of lifeguard data in New Zealand between 2001-2022 comes from intermediate (bar-rip) beaches, where 86% of all lifeguard patrol hours are undertaken, while 10% and 4% are undertaken at dissipative and reflective beaches, respectively. Furthermore, 96% of rip current rescues and 95% of the highest-hazard events (P_{rip} > 90$^{\mathrm{th}}$%ile) occurred at intermediate beaches. Dissipative and reflective beaches in comparison contribute far fewer rip incidents, at 3% and 1%, respectively, and fewer high-hazard events, at 5% and 1%, respectively. It is acknowledged that because the vast majority of lifeguards operate on intermediate beaches there is data imbalance between the three beach types.'

Comment (23)

- Table 1 – it occurs to me that the 'type' of rip current has not been described enough (or the lack of reporting of the rip type).

As per our reply to your comment 20 above, rip current type was not included in table 1 as it was not measured during the study or reported by lifeguards. We describe rip current types briefly in the introduction in relation to the types of forcing that relate to different rip types and refer to them briefly in the conceptual figure in the discussion section.

Comment (24)

- Suggest Table 2 is not needed and can be incorporated within text

Thanks for this suggestion. I've now removed table 2 and made sure the values are all captured in the text instead. Lines 396-400 now say:

> 'The efficiency of the algorithm (Eq. (5)) at predicting incidents increases as the RI level increases. Looking at all warnings in excess of RI 1 the efficiency is 1.06, meaning that a slightly larger proportion of incidents happen at RI 2-5 than the proportion of time those warnings are given. However, for RI > 2 efficiency increases to 1.23, and at RI > 3 efficiency is 1.63. This suggests that warnings above those thresholds are given sparingly relative to the proportion of incidents they capture. For example, RI 4 was predicted 20% of the time but captured 42% of all SLSNZ rip incidents.'

And L409-412 now say:

> 'As with rip current risk, the algorithm is increasingly efficient at capturing high-hazard events as RI increases. For warnings above RI 1, the efficiency is 1.05 meaning the algorithm captures a slightly larger proportion of high-hazard events than the proportion of time warnings are given. The efficiency increases to 1.22 at RI > 2 and to 1.66 at RI > 3. This means that high-hazard events occur frequently at RI 3-5 relative to the number of warnings given.'