ARTICLE IN PRESS

Journal of Pyrogeography xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Pyrogeography

journal homepage: www.sciencedirect.com/journal/journal-of-pyrogeography

Fifty shades of "great": User-informed principles for wildfire simulator development and utilisation

Caitlin Symon [a,c,*], Timothy Neale [b,c], Gabrielle Miller [b,c], Alexander I. Filkov [a,c], Kate A. Parkins [a,c], Erica Marshall [a,c], Trent D. Penman [a,c], Hamish Clarke [a,c]

[a] FLARE Wildfire Research, School of Agriculture Food and Ecosystem Sciences, The University of Melbourne, Grattan St, Parkville, Victoria, 3010, Australia
[b] Deakin University, Victoria, 3125, Australia
[c] Natural Hazards Research Australia, RMIT University, Building 8, 360 Swanston St, Melbourne, VIC 3000, Australia

ARTICLE INFO

Keywords:
Fire behaviour
Wildfire
Fire management
Modelling
Simulation

ABSTRACT

Fire simulators play a significant role in contemporary fire management. Their importance is likely to grow as fire risk shifts in response to climate change, land use change and other drivers. Despite their widespread use, there are major gaps in our knowledge about how fire simulators are used and what is required to ensure they support management and research. To guide the future development and use of fire simulators, we undertook a comprehensive engagement process with simulator users across Australia. This involved a staged, mixed-methods approach consisting of interviews, workshops, and a questionnaire to identify important technical and contextual factors relevant to simulator development and use. The findings were consolidated into a defined set of simulator performance criteria. Numerical estimates were obtained for several technical criteria but there was high uncertainty around these values. Contextual criteria, while viewed as important, proved difficult to benchmark and were not suited to numerical estimates. We argue that simulator development guided by a set of principles, rather than set-and-forget standards, will better accommodate diverse user requirements, including the importance of non-technical factors and the rapid pace of change in simulator technology. The qualitative and quantitative findings from participant engagement were used to derive four principles for future simulator development. These principles are 1) Drive simulator performance through improved modelling and data, 2) Improve usability of fire simulator software and hardware, 3) Adopt a comprehensive and transparent approach to validation and verification, 4) Maintain a cohesive approach to development and use through governance, capacity building and engagement.

1. Introduction

Increases in likelihood, severity, and impact of wildfires under climate change are intensifying the pressure on emergency services and land management agencies to respond to and mitigate fire risks (Bowman et al., 2020; Filkov et al., 2020; Jones et al., 2022). In recent decades, the Australian wildfire sector has come to value fire simulators as a tool for strengthening shared knowledge of landscape fire and to support decision-making around their management (Slijepcevic et al., 2008; Sullivan, 2009). This sector includes management, research, and emergency response agencies and organisations that contribute to the development and implementation of timely and locally appropriate fire management practices. Fire simulators are computer programs used to depict processes involved in the behaviour and spread of fire through a landscape. These tools are built upon physical and empirical models of fire behaviour that require inputs such as terrain, weather, vegetation, and assets (Sullivan, 2009). Fire simulators allow users to make predictions and explore scenarios that would be impractical to test in the field due to constraints around cost, resources, safety and time (Sullivan, 2009; Cruz et al., 2014). Critical uses of fire simulators include tactical response to ongoing fire, strategic planning of fuel management, and research into fire behaviour and risk. We use the terms 'fire simulators' and 'simulators' interchangeably throughout this paper, but other terminology includes fire behaviour simulators, wildfire simulators, automated fire prediction, operational fire modelling, and fire spread models embedded in decision-support tools (however, we exclude virtual and

* Corresponding author. FLARE Wildfire Research, School of Agriculture Food and Ecosystem Sciences, The University of Melbourne, Grattan St, Parkville, Victoria, 3010, Australia.
E-mail address: caitlin.symon@unimelb.edu.au (C. Symon).

https://doi.org/10.1016/j.pyro.2025.100001
Received 15 July 2025; Received in revised form 23 October 2025; Accepted 28 October 2025
Available online 4 November 2025

Please cite this article as: Caitlin Symon et al., *Journal of Pyrogeography,* https://doi.org/10.1016/j.pyro.2025.100001

augmented reality simulation). Given the significant threat that wildfires pose to communities, ecosystems and emergency service staff, and the limited resources available for simulator development, it is crucial that the technologies employed by this sector are appropriate and effective.

Since fire simulators were first introduced in the 1970s and 1980s, there has been pressure both for their uptake and continual improvement in agencies and research institutions (Van Wagner, 1985; Neale & May 2018). The number of simulators in use and the variety of their usages has grown over recent decades (Pacheco et al., 2015; Parisien et al., 2019; Begg et al., 2021). This is reflective of a broader reliance upon new technology to support risk assessment, prediction, and decision-making (Jasanoff, 2005; Chen et al., 2019). Whilst this reliance is not a problem per se, the assumption that new technology is inherently good, and better than its predecessors, can overlook an important consideration of whether the new option genuinely improves outcomes (Borup et al., 2006).

To date, fire simulator research and innovation have largely focused on the technical performance of simulating tools, particularly on improving output accuracy (Opperman et al., 2006; Fox-Hughes et al., 2024). However, the importance of additional factors outside of the core fire behaviour models has long been recognised (Van Wagner, 1971; Albini, 1976). As important as technical performance and capability are to establishing high-quality and accurate fire simulators, there is an intractable 'social dimension' in the development and application of simulators (Neale & May 2018). This relates to social, organisational, and structural factors, including the infrastructure, governance, support, resourcing, and capabilities required to operate fire simulators. These factors are hereafter referred to as 'contextual factors', which, in combination with technical ones, collectively influence whether a simulator is useful and effective (Neale, 2016; McFayden et al., 2024). Research on wildfire decision-support systems that incorporate fire simulators similarly emphasises the importance of contextual factors in the effective design and use of those tools (Pacheco et al., 2015; Rapp et al., 2020; Calkin et al., 2021; Colavito, 2021). There relevance is not unique to fire simulators and is evident in the development of early warning systems (Garcia and Fearnley, 2012), environmental modelling (Parker, 2020; Beven and Lane, 2022; Hamilton et al., 2022), human-tool relationships (Bainbridge, 1983; Onnasch et al., 2014; Strauch, 2017), and human-computer interactions (Carver and Turoff, 2007).

Fire managers have repeatedly acknowledged the importance of contextual factors that, whilst external to the tool itself, can significantly impact a simulator's fitness-for-purpose (Alexander and Cruz, 2013b; Cruz et al., 2014; Neale et al., 2021). Nevertheless, the influence of these factors on simulator outcomes is still poorly understood and enhancements to technical factors remains the priority (Cruz et al., 2014).

There is a significant knowledge gap regarding how to strategically balance technical and contextual factors specific to fire simulator development and use (Neale et al., 2021). Building a qualitative understanding of stakeholders' needs and perspectives is an essential step for addressing this gap. Our study engaged with simulator users to identify both the technical and contextual factors that collectively determine whether a fire simulator is seen as 'good[1]' by users. The research aimed to.

a. Understand current user requirements and priority factors relevant to future development for the broad and expanding community of fire simulator users.
b. Establish a set of standards or measurements that reflect simulator requirements to guide future simulator development.

[1] The term 'good simulator' is used throughout the text. This wording has been intentionally chosen as it allows space for users to define what is meant by 'good'.

2. Methodology

This study aims to inform future simulator development with the findings. As such, the scope covers fire simulators collectively, rather than any individual simulator. A broad perspective is essential given both the diversity of simulators currently in use and the regular appearance of new simulators. Nevertheless, it may help readers to note that fire simulators can be broadly categorised into three main types based on their underlying fire spread models: physical and quasi-physical models, empirical and quasi-empirical models, and coupled fire-atmosphere models. Each type offers different capabilities and trade-offs between factors related to simulator performance. The users in this study predominantly work with simulators employing empirical and quasi-empirical fire behaviour models, which provide faster than real-time fire spread predictions and are common in operational settings.

2.1. Participant selection

We focused on simulator users in Australia i.e. individuals who actively engage with fire simulator tools and are responsible for developing simulator outputs for audiences. Initial project development took place in collaboration with Australasian Fire and Emergency Service Authorities Council (AFAC) Predictive Services Working Group, the Country Fire Authority (CFA) and the Rural Fire Service (RFS). AFAC are the peak body for fire, land management and emergency services in Australia and New Zealand. The CFA and RFS are state-based firefighting organisations in Victoria and New South Wales, respectively.

Early consultation with AFAC, CFA and RFS led to the establishment of three simulator use cases that broadly categorised different types of fire simulator users: Tactical, Strategic, and Research. Tactical and Strategic users are both found in the fire management and operations sector. Tactical use focuses on assessments of current or potential fire behaviour and spread in the short-term future, and Strategic use relates to longer-term planning and other non-tactical operational use. Research, on the other hand, mostly takes place in academic institutions and tends to cover a broad range of applications. A fourth use case (Other) was created to capture additional users that did not fit within any of the three primary categories, such as training and education uses. This study aimed to represent a broad cross-section of the user community, spanning jurisdiction, sector and use case (Table 1). The term

Table 1
Summary of study participants across each phase of engagement.

	Interviews	Workshop 1	Workshop 2	Questionnaire
Jurisdiction				
Australian Capital Territory	2	3	1	2
New South Wales	4	4	5	17
Northern Territory	0	0	1	2
Queenslan	1	3	3	8
South Australia	2	2	4	2
Tasmani	1	1	1	6
Victoria	3	6	3	22
Western Australia	1	0	2	5
National	4	7	5	3
Overseas	3	0	0	1
Use Case				
Tactical	7[a]	7	9	37
Strategic	8	10	11	18
Research	6	7	4	9
Other	1	2	1	3
Gender identity				
Female	3	7	8	13
Male	18	19	17	53
Prefer not to say	0	0	0	1
Total	**21**	**26**	**25**	**67**

[a] One participant identified as a Tactical and Strategic user.

'jurisdiction' refers to the geographical area, defined as an Australian State or Territory, in which a participant's work is carried out. Consultation with key end users also informed participant selection, which was supported by snowball[2] sampling. In each phase of engagement, participants responded to prompts according to the most relevant simulator use case given their expertise and experience. Participants' answers were analysed individually and with respect to their self-assigned use case.

2.2. Stakeholder engagement

This study used a phased, mixed-methods approach that combined responses from interviews, workshops and a questionnaire, which provided stakeholders the opportunity to contribute across multiple means of discussion and engagement (Fig. 1).

The interviews were conducted to identify priority technical and contextual factors relevant to simulator development and use among users and to identify themes among participant responses. During initial project development and throughout engagement, numerous, often interrelated, technical and contextual factors were put forward. For clarity, these factors were refined into distinct categories, which form the set of performance criteria, hereafter referred to as 'criteria', relevant to simulator development and use. These criteria were refined throughout all phases of stakeholder engagement in response to participant feedback. They also formed the basis of the two workshops and Part 1 of the questionnaire. The two workshops were designed to refine and rank criteria and test their potential for benchmarking. The workshops and questionnaire, whilst obtaining qualitative information from participants, also elicited quantitative data relating to user standards for the criteria.

The design of the questionnaire, informed by responses from the interviews and workshops, was intended to deepen the representation of simulator users and to identify patterns and consistencies in participant's expectations around simulator performance and development. It also tested whether factors related to simulator development and use, including a subset of the established simulator performance criteria, could be quantified and whether such measures would provide useful insights.

Based on the sum of engagement, we derived four principles for simulator development and use. These principles were designed to encapsulate the range of participant comments and contributions in a way that can guide future actions.

2.3. Interviews

Semi-structured interviews were conducted online with 20 subject matter experts. The interviews were guided by a schedule of questions designed to discuss the technical and contextual factors relevant to simulator development and use and identify criteria and themes. The interviews engaged with topics such as: who is using simulators and for what purposes?; limitations, knowledge and data gaps within existing simulators and their use; factors relevant to simulator performance; and treatment of uncertainty and future challenges in simulator development and use (Supplementary Table 1). The audio from each interview was recorded, transcribed and analysed using the qualitative analysis software NVivo 10. The findings from these interviews also informed the design of the workshops and the questionnaire. Interviewees were categorised into the three use cases (Tactical Operations [TO; n = 7], Strategic planning [SP; n = 8], Research [Res; n = 6]), with 'n' denoting the sample size of participants in each category. One participant was categorised as Other [O; n = 1]). To protect participant anonymity, quotes from interviews have been coded to reveal the jurisdiction and use case of each participant only (e.g., "SA-Res1" for South Australian researcher #1, "Vic SP1" for Victorian Strategic planning #1, "WA-TO1" for Western Australian Tactical Operations #1, etc.).

2.4. Elicitation workshops

Two online elicitation workshops were held (n = 26, n = 25). Participants included interviewees as well as other subject matter experts who represented a diversity of agencies and jurisdictions across Australia.

Simulator use and development priorities identified during early consultation with key stakeholder groups (AFAC, RFS, CFA) were combined with findings from the interviews, resulting in a set of 14 criteria for 'good' simulators (Table 2). Because some criteria could be interpreted in multiple ways (e.g. speed), the initial list of 14 criteria was expanded to 35 (Supplementary Table 2). These initial sets of criteria were introduced to participants in the workshops to explore their willingness and ability to prioritise and parameterise what makes a good fire simulator, as well as their willingness to adopt minor and major simulator changes relating to these criteria.

In the first workshop (Supplementary Material 3), participants were asked to rate the importance of each of the 35 criteria (Supplementary Table 2) from 1 to 5, where 1 = 'not important', 3 = 'moderately important' and 5 = 'very important'. The workshop then focused on three hypothetical scenarios. Participants were asked to determine the amount of change they would need to observe in each of the 14 criterion (Table 2) for them to adopt a) a new simulator data layer, b) a new simulator version, or b) a new simulator altogether. Participants were also asked about the circumstances in which they would make their decision unhesitatingly; they would consider adoption; and they would reject the adoption outright.

The second workshop considered whether simulator criteria (Table 2) could be associated with generalisable performance expectations or instead were context dependent (Supplementary Material 4). Participants were asked to, where possible, identify objective, numerical benchmarks for their most important criteria, for three levels of performance: unacceptable performance; performance that would justify minor or major simulator changes; and ideal or 'game-changing' performance. When quantitative benchmarks for certain variables were said to be either inapparent or inappropriate, participants were asked to provide qualitative expectations for their performance instead. Participant contributions from both workshops helped to refine the simulator performance criteria and also informed the development of the questionnaire. Participant responses were recorded using Qualtrics and by note takers.

2.5. Questionnaire

The final phase of engagement was an online questionnaire made available to fire simulator users across Australia. It was open to all users regardless of skill and experience. 67 individuals completed the questionnaire. Snowball sampling was used to identify and engage simulator users throughout Australia. The questionnaire consisted of 25 questions (see Supplementary Material 5). Part 1 of the questionnaire aimed to test whether generalisable standards, in the form of measurable (numerical) benchmarks, could be found for four criteria that had been identified in the workshops as potentially quantifiable. All four were related to technical factors. They were accuracy, robustness of modelling framework (capacity to run ensembles), scale, and speed. Note that accuracy was considered to be a measurable component of the validation criterion (Table 2). 'Accuracy' was considered with respect to four fire characterisation variables (forward rate of spread, area burned, fire behaviour, and impacts), 'Scale' referred to both spatial resolution and time-step

[2] A non-probability sampling technique that uses current participants to assist in the recruitment of new participants, in this instance, fellow colleagues.

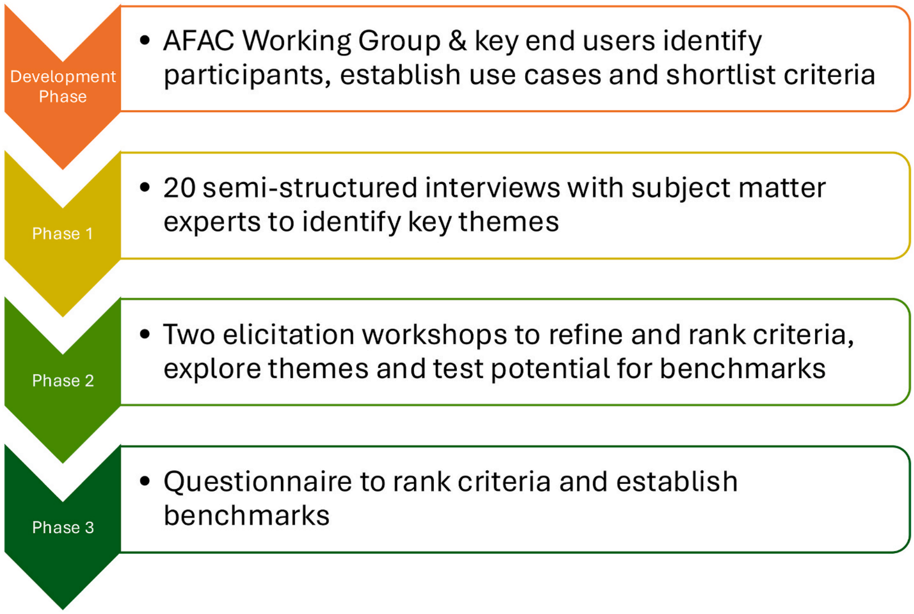

Fig. 1. The study began with a development phase and proceeded through three phases of engagement, with each phase of engagement informing the next.

Table 2
Initial set of simulator performance criteria, which formed the basis of subsequent consultation.

Criteria	Description
Configurability	Users can easily adjust models, inputs, and outputs
Compatibility	Simulator is compatible with other models, agency systems and policies
Ease of use	Intuitive interface; Reliable under diverse conditions
Effectiveness of software framework	Functions on multiple platforms and offline; Open source
Handling of inputs	Easy to prepare inputs; Results are sensitive to inputs
Handling of outputs	Outputs are clear, tailored to audiences, easy to store and audit
Robustness of modelling framework	Easily handles ensembles and uncertainty; Results are traceable
Scale	Appropriate spatial and temporal resolution
Speed	Quick fire behaviour simulation, quick inputs, quick outputs
Support	Training, documentation and support to run simulator and interpret output
Trustworthiness	Users and agencies trust the simulator, developers and underlying research
Validation	Simulators and underlying models are accurate and validated
Value for money	Simulator investments provide value for money
Versatility	Simulator handles all aspects of fire behaviour (Some prefer specialists)

resolution, ‘Speed’ referred to output generation time, and ‘Robustness of modelling framework’ referred to the number of ensembles that users required. Users were asked to identify minimum acceptable[3] and maximum necessary[4] thresholds for performance for these criteria. Part 2 of the questionnaire aimed to identify consistencies in factors important for simulator users across several criteria that were harder for participants to provide measurable benchmarks for. They were derived from the criteria initially identified in the interviews and from the workshop discussions. All five were linked to contextual factors. They were user experience, output quality, outcome quality, trustworthiness and the broader ‘simulator ecosystem’. User experience reflects the extent to which a tool aligns with a user’s abilities and requirements. Output quality refers to the technical appearance of what is produced from a fire simulator, while outcome quality refers to how well that output improves decision-making, actions, results, and consequences from the use of simulator outputs. Trustworthiness reflects the degree of support for, and acceptance of a simulator’s outputs. The simulator ecosystem is a term that emerged out of early participant engagement. It captures the broader context and systems relevant to the use of simulators. It includes the frameworks, governance, support systems, resources, and stakeholders that collectively influence outcomes from a simulator’s implementation and use. For each criteria, participants were presented with 10–14 factors associated with that criteria. These factors, both technical and contextual, were derived from participant responses in both the interviews and workshops. Participants were asked to select and rank the top 50 percent of those factors according to their value to the participant. Participants were also given the opportunity to provide qualitative input about what they believe makes a good fire simulator. In the results, we highlight the most popular factors relevant to each of the criteria.

3. Results

3.1. Interviews – emergent themes

Several themes emerged throughout the interviews. These were collaboration and coordination; reputation, trust, and confidence; interpretation and communication of outputs; configurability and automation of simulators; user prioritisation; and ease of use (Table 2). These themes consistently appeared in the interviews, although participants’ views on how they should be addressed varied significantly. This was perhaps best encapsulated through the common although ambiguous use of the term ‘fit-for-purpose’ to describe good simulators.[5] Fitness-for-purpose is a general term designed to capture whether a tool is appropriate in supporting users to achieve outcomes relevant to their

[3] The amount, below which simulator performance would be unacceptably low for use by a participant.

[4] The amount, above which any increase in performance would have a negligible effect on the quality of a participant’s output.

[5] The term ‘good simulator’ is used throughout the text. This wording has been intentionally chosen as it allows space for users to define what is meant by ‘good’.

role requirements and responsibilities. Although there was consensus among participants that a good simulator meant one that is fit-for-purpose, perspectives on what this entails were not generalisable because the purposes for simulator use were equally diverse. A simulator's purpose depended upon the circumstances of the user and the context of use and could even vary for a single user depending upon what task they were undertaking or what *"hat they're wearing"* (NSW-SP2), with many participants covering multiple roles and responsibilities within each role. As one participant stated, *"it's about a simulator that's useful for a decision-maker at the time"* (Vic-TO1).

Good fire simulators were seen to rely upon the effective functioning of a web of factors or, as some described it, the simulator ecosystem (Fig. 2). For some users, the most important factors pertained directly to the performance of a simulator, however, often these factors related to the broader structures and systems operating outside of the physical tools. As one participant stated, "*It's not just having a simulator that does stuff. It's the ecosystem, if you like, around it and the context for how we're building simulators, which is really important*" (Aus-Res2). These technical and contextual factors were highly interconnected. Decisions on how to improve one aspect of a simulator were suggested to have wide-reaching implications and the prioritisation of one factor frequently overlapped with, and sometimes conflicted with, efforts improve others.

An illustration of this interconnectedness between factors was the debate around who should have access to simulators (i.e. experts only or open access). This was a divisive topic among participants and crossed over multiple themes such as ease of use, user prioritisation (e.g. expert vs general user), and configurability vs automation. It also had implications for contextual factors, such as training and support, trust and confidence, and responsibility and accountability for outputs. Although participants consistently mentioned factors related to technical performance, contextual factors dominated their responses, with many participants focused on the broader structures required to support effective simulator use.

3.2. Workshops - Refinement of criteria and potential for benchmarks

All simulator criteria presented to participants (Supplementary Table 2) were viewed as important by workshop participants, with little consensus as to their relative importance. When asked to provide a rating out of 5, the average rating for each of the criteria was 4.0 (standard deviation 0.30). In both workshops, users were also generally unwilling or unable to provide discrete, quantitative thresholds or benchmarks for performance of each of the criteria (Table 2). The responses they put forward did not appear to provide a useful basis to inform management decisions due to (a) high variability and a lack of consistency between participants' responses, and (b) a lack of confidence among participants in their answers, with many wanting to qualify their choices in the group discussions. Table 3 provides common reasons for this lack of confidence in quantifiable benchmarks. These same reasons also meant that, when asked to quantify the amount of change to simulator software needed to justify the effort of adoption, respondents did not agree on any concrete or absolute value. Instead, participants preferred to discuss qualitative or contextual factors that determine whether a simulator is good as well as what might influence their willingness to accept any changes to the simulator they use.

3.3. Questionnaire - Ranking criteria and establishing benchmarks

3.3.1. Criteria benchmarks identified by questionnaire participants

3.3.1.1. Accuracy benchmarks. User expectations for a simulator's accuracy at characterising rate of spread, area burned, fire behaviour,[6] and impacts[7] were largely consistent across users and each of the four measures (Fig. 3). The median estimate for minimum accuracy required for participants, irrespective of use case (Tactical, Strategic and Research) was 60 percent, while the median maximum accuracy required ranged from 85 to 90 percent across the four measures. The median accuracy required was consistently lower for Research users than Tactical and Strategic, however, the differences were marginal.

3.3.1.2. Performance benchmarks. Large ranges were observed for each of the four performance variables: spatial resolution, time-step, ensembles and speed (Table 4). This was true for minimum required performance and maximum necessary performance. As such, none of the estimated minimum and maximum values for any of the four variables displayed definitive performance thresholds that are likely to be useful for informing simulator design and development. These large ranges were consistent across Tactical, Strategic and Research users which, combined with the similarities in median minimum and maximum values across the three use cases, indicates that use case is not a clear distinguisher of expectations for these four performance variables (Fig. 4)

3.3.2. Important factors related to simulator use and development

When questioned about user experience, output quality, outcome quality, trustworthiness, and the broader simulator ecosystem, a range of factors were deemed important and there were few standouts. Participants' most important factors also varied between technical and contextual ones, some of which directly related to a simulator's performance, and others which broadly related to the simulator ecosystem. Whilst some variability was evident between the three use cases, Tactical, Strategic and Research, the priority factors of different users were largely consistent across participants and differences observed between use cases were minor.

To indicate user preferences, the most popular factors within each were identified based on how often they appeared in participants' top 50 % ranked variables.

- For user experience, stable simulator platform (94 %), ability to adjust simulator settings (88 %), intuitive graphical user interface (GUI) (75 %) were the most popular factors. The importance of a 'Stable simulator platform' was reflected in several participant comments. They stated that, for a simulator to be good, they needed the "ability to rely on it". It benefited from a "standalone operation", that is "not reliant on connectivity", and "with redundancy, for example, to cope with local power/network outages, operate locally in remote (no network) locations" (various questionnaire participants).
- For output quality, the three standout factors all related to data. These were accurate fuel input data (88 %), accurate weather input data (87 %), and accurate mapped location (67 %). Participant comments on data also mentioned the need for "enhancements to data retention and validation" and "training data sets" (various questionnaire participants). Several participants also stated the need to better capture and "integrate real world information/data" from various sources (various questionnaire participants).

[6] Fire behaviour relates to the characterisation of factors such as intensity, flame height, and ember transfer.

[7] Impacts relate to the characterisation of consequences for values such as communities, the environment, and asset loss.

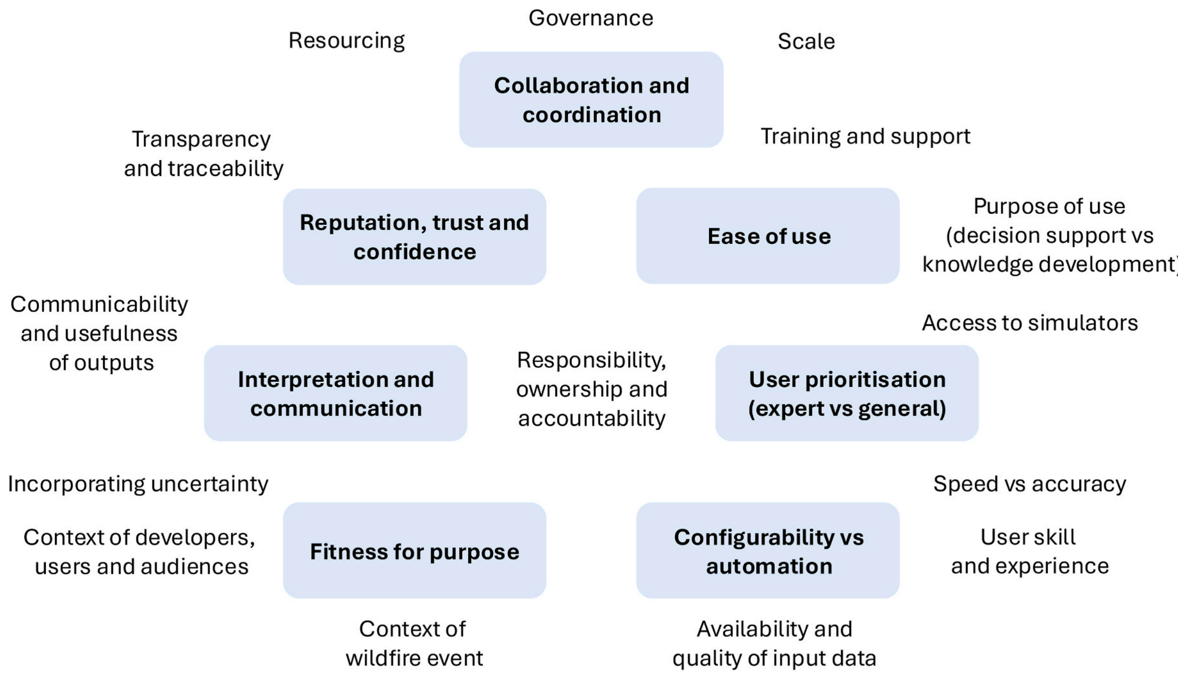

Fig. 2. This figure displays the reoccurring themes (boxed) and factors (unboxed) for simulator development and use that emerged from the semi-structured interviews with fire simulator users. This figure serves as a visual representation of the findings from the interviews and is not meant to be analysed as an absolute representation of the connection between all themes and relevant factors associated with simulator development and use.

- Priority factors for outcome quality were more balanced across several options. The most popular were clear communication of assumptions (69 %), quality of the visualisation (61.2 %), simplicity of the overall output presentation document (58.2 %).
- For user trustworthiness, the most popular factors were accuracy of outputs (75 %), validation of simulator (70 %), Clarity around how outputs are produced (assumptions and limitations)" (57 %), and peer-reviewed science (55 %). As one questionnaire participant stated, "In the end, a good simulator is one that gives accurate predictions that can inform timely decision-making".
- Lastly, simulator ecosystem priority factors at the (a) organisational and (b) national level had the same top priority factors but with variation in lower order factors.

a. Important development factors at the organisational level were, improved fire behaviour models (66 %) and improved validation of data, models, and simulators (64 %). These were followed by compatibility with existing mapping (52 %), training for users (51 %), and compatibility with existing systems (51 %).

b. National level priority factors were, improved validation of data, models, and simulators (66 %) and improved fire behaviour models (60 %). These were followed by improved transparency & documentation (58 %), improved governance (52 %), and improved inputs - data collection & storage (49 %).

4. Discussion

Our study explored the technical and contextual factors that collectively support the creation fire simulators. Through engagement with fire simulator users in Australia via workshops, interviews and a questionnaire our research addressed our first aim of shedding light on current requirements and priorities for future development for the broad and expanding community of fire simulator users. We found that grouping simulator users into use cases and refining technical and contextual factors into simulator performance criteria (Table 2, Supplementary Table 2) was useful for structuring discussions. However, this approach was less effective in articulating distinct user requirements. Consequently, our second aim - pinpointing and codifying user requirements into concrete thresholds - proved challenging. The findings suggest that establishing thresholds may be an ineffective approach for supporting the development of fire simulators that users recognise as 'good'.

Through the phased engagement a set of criteria emerged that were important for simulator performance: ease of use, speed, configurability, versatility, robustness of modelling framework, effectiveness of software framework, handling of inputs, handling of outputs, scale, validation, support, trustworthiness, compatibility, and value for money. However, user requirements for each of these criteria were diverse, context dependent, evolving and resistant to simplification. Three major use cases were also identified during early discussions with project end-users: Tactical, Strategic and Research. Despite the conceptual clarity of these categories, clear differences in the needs and priorities between the three use cases were the exception rather than the rule (see Figs. 3 and 4). Potential reasons include differences in user skill and experience level and organisational or jurisdictional factors. One questionnaire participant outlined some of the complexities involved:

> "A simulator may be considered "good" and highly applicable to one fire but "not good" and completely unusable for another fire due to poor input information, deficiencies in the fire behaviour models etc. A poor user may not be able to adequately identify when a simulator is good or poor for the circumstance and could inappropriately use a good simulator resulting in poor outputs. Some geographic areas may always be inappropriate for simulator use due to failure to characterise how fires truly burn there."

Progress towards objective standards may be possible for a limited set of technical variables such as accuracy, spatial resolution, model timestep, output speed and number of ensembles. Yet even within these categories users varied strongly in their definitions and requirements, suggesting a challenge for any benchmarks in capturing the diversity of user needs without being impractically broad or detached from realistic expectations of performance. For example, some users in this study indicated an expected accuracy of up to 100 percent, which is challenging for current models (Fox-Hughes et al., 2024).

Throughout the interviews and workshop discussions, participants

Table 3
Common participant reasons for a lack of quantifiable criteria benchmarks.

Reason	Details
Context dependency of simulator use	Context dependency made the value of specific criteria hard to quantify or parameterise. 'Good' fire simulators varied according to (a) the broader circumstances of the user (e.g. their organisation, jurisdiction, use case, experience level) and (b) the context of any given use (e.g. type, size and risk of fire, purpose of the simulation, audience). Participants were able to make broad statements about how the presence or absence of certain features or capabilities would affect the quality of a simulator or their desire to change simulators. They were less able to pinpoint specific expectations for benchmarks that reflected their own needs, let alone to generalise about the needs of the broader simulator community.
Specific, measurable values for user expectations were challenging to conceptualise	Participants resisted prompts to articulate tangible benchmarks that accurately encapsulated their requirements for a given criteria. The exceptions were some technical performance criteria (e.g. accuracy, speed, resolution, and number of ensembles), for which a small subset of participants were able to provide numeric values.
Interconnectedness of criteria	Criteria were said to be interconnected and therefore not easily evaluated in isolation or traded off against each other.
Importance of contextual factors	Contexual factors were considered equally, if not more, important as technical factors. Contextual factors also were typically the hardest to evaluate and quantify.

frequently emphasised the importance of contextual factors, often placing them above technical factors. Their expectations and benchmarks for contextual factors, however, were consistently the hardest to articulate and measure, often proving intangible, interdependent and connected to other social, logistical and environmental challenges of land management and research more broadly. One workshop participant outlined the importance of contextual factors:

> "You can have a perfect model or a perfect simulator, and if you don't have people making good decisions based on it, then it's as good as useless … It's as much about who's using it and how they're using the information that comes out of the back end of it."

These findings suggest that quantified benchmarks or standards may not be the most suitable way to ensure fire simulators are fit-for-purpose for all users. Instead, fire simulators should ideally be responsive to the needs of different users, audiences and use contexts, and adaptable to changes in simulator use over time (McFayden et al., 2024). "Good" simulators require a capacity to integrate new science and knowledge, adopt technological advances, and respond to changing user requirements. Despite the appeal of using quantitative benchmarks to ensure quality control in the development of fire simulators (Fox-Hughes et al., 2024), for this to work, the guidelines that dictate their development and use may need to be as flexible and as adaptable as the tools themselves. Flexibility and adaptability have previously been identified as important features of new simulating platforms (Miller et al., 2015). Our research suggests this must extend beyond the tools themselves. Common examples put forward by participants related to data quality and access; support systems, staff capacity, funding and resources; infrastructure such as internet access, reception, and data storage; organisational structure, policy and culture; and the varied needs of audiences and other stakeholder groups. All these examples require coordination and effort from not only simulator development teams but also the broader Australian wildfire sector. This, in turn, requires consistent engagement and communication between stakeholders, fortified by the proven and enduring reliability of the tool.

To achieve this, the wildfire sector would benefit from methods for systematically integrating diverse technical and contextual factors into simulator development that are grounded in the local context (McIntosh et al., 2011; Merritt et al., 2017; Walling and Vaneeckhaute, 2020). This includes centring the knowledge and perspectives of stakeholders from across the fire simulator community during research and development. Beyond the research and development stages, our research drew attention to the fact that the effective use of new or updated tools requires support and capacity building during implementation and throughout the lifetime of the tool. Additionally, these supports must be adapted to specific user and audience requirements. Such requirements for development, implementation and use of simulators are unlikely to be well-represented with fixed, numerical performance benchmarks. To support the future use and development of fire simulators, we identified four principles to address performance needs in a way that tailors development to different use cases, jurisdictions and sectors (Table 5). These principles, their components and actions to consider represent a synthesis of the quantitative and qualitative results obtained across the three phases of engagement. We briefly discuss each of these principles below.

4.1. Drive simulator performance through improved modelling and data

Improvements to simulator capability and performance are integral to the quality of simulator outputs and have driven much improvement already. Study participants put forward a range of improvements relating, firstly, to underlying fire behaviour models and, secondly, the availability and quality of input data. Both were identified as a major limitation to the accuracy, usefulness and evaluation of simulator outputs, echoing previous research (Sullivan, 2009; Alexander and Cruz, 2013b; Cruz et al., 2014). The list of suggestions for improvements to modelling and data was extensive and opinions about their prioritisation were strongly varied and user dependent. Continual improvements to fire simulator performance require an awareness of and responsiveness to the current science and technology, but this alone will not be enough to create good simulators. This will depend upon a sustained effort to engage with users' suggestions, understand their diverse and changing priorities, to discern how changes may impact users differently, and to design suitable trade-offs between options. Additionally, coordinated efforts between those collecting, managing and using the data are required and improvements in this area will also inform new capacities such as machine learning in wildfire simulations (Jain et al., 2020; Ghali and Akhloufi, 2023).

4.2. Improve usability of fire simulator software and hardware

Simulator usability incorporates simulators, their outputs and broader simulator infrastructure. User-centric design was seen as essential by participants. As individual fire simulators are typically used by many users, to be fit-for-purpose they must accommodate a variety of different types of users, audiences and use contexts. Users consistently stated a desire for simulators to be modular where possible, allowing for specific features to be switched on or off depending upon a user's skills and requirements. The importance of modularity in place of blanket standardisation has similarly been discussed in previous research on simulator development (Miller et al., 2015), for situational awareness during wildfire response operations (Goubran et al., 2016) and in emergency warning systems (Fearnley et al., 2012).

The communicability of a simulator output is another important

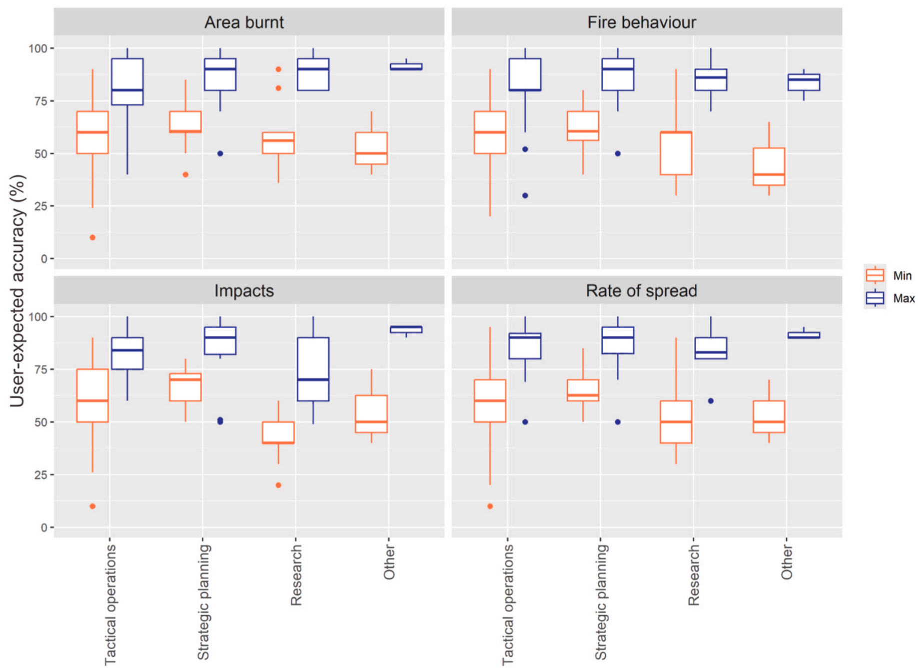

Fig. 3. Performance benchmarks for four aspects of fire characterisation: area burnt, fire behaviour, impacts, and rate of spread. Orange boxplots show minimum acceptable threshold (the amount, below which simulator performance would be unacceptably low for a participant's use), blue boxplots show the maximum necessary threshold (the amount, above which any increase in performance would have a negligible effect on the quality of a participant's output). Results are separated according to the four simulator use cases. The lower and upper hinges correspond to the first and third quartiles. Whiskers extend up to 1.5 x the interquartile range.

Table 4
Benchmarks for performance of simulator features; spatial resolution, model timestep, output speed, and number of ensembles (minimum required and maximum necessary). Note: Timestep was intended to capture the internal timestep i.e. the increment at which the model advances. It is possible that some respondents were referring the timestep of potential outputs. Results should be interpreted with this caveat in mind.

Variable	Minimum required (median)	Maximum required (median)	Minimum required (range)	Maximum required (range)
Spatial resolution	30m^2	200m2	1–30m^2	30 - 10,000m^2
Timestep	10 min	60 min	1 s–3 h	1 min- 24 h
Output speed	1 min	15 min	1 s–5 h	30 s - 2 weeks
Number of ensembles	1	50	0–20,000 ensembles	3–20 million ensembles

factor of their usability. It requires balancing the provision of enough information to support informed decision-making without confusing or overwhelming audiences. Outputs need to reflect audiences and their circumstances (Garcia and Fearnley, 2012; Cheong et al., 2016; Morrison et al., 2024) and may therefore be resistant to standardisation (Scolobig et al., 2022). Additional factors relevant to usability can affect the integration and uptake of a simulator. They include the hardware and infrastructure required to operate a simulator, platform stability (particularly for rural and remote users), and reliable support (McFayden et al., 2024).

4.3. Adopt a comprehensive and transparent approach to validation and verification

Validation and verification of fire simulators and their inputs was said to be important to participants but challenging to do well. Doing so improves accuracy and is also fundamental to trust and support from both users and their audiences (Alexander and Cruz, 2013a; Fox-Hughes et al., 2024). Although there has been much progress in methods for evaluation (e.g. Kelso et al., 2015; Duff et al., 2018), this progress is undermined by the absence of robust mechanisms for reconstructing wildfire events and challenges in obtaining suitable wildfire data (Filkov et al., 2018; Fox-Hughes et al., 2024). Improvements in this space will require significant investment and collaboration between researchers, simulator developers, and fire agencies (Alexander and Cruz, 2013a).

Participants also emphasised the value in transparent and traceable documentation of simulator use processes to trace significant drivers of outputs, such as information about how a simulator was used during the generation of an output - including input data, assumptions, run time, issues or errors - and the specific details of the output. It was suggested that documenting this information will build transparency around why, in a specific context, a simulator produced a particular result. Making this information available to both the user and, when necessary, their audiences can support individuals in making fast, confident and well-informed decisions and may also be valuable in informing improvements to future simulators (Plucinski et al., 2017). Simulator users also indicated that the evaluation of simulators should be expanded to include the fire management-related outcomes of simulator use, similar to the evaluation of wildfire decision-support systems (e.g. Pacheco et al., 2015; Rapp et al., 2020). As one participant stated,

> "We need to focus on the outcomes, not the outputs … What are the decisions being made as a result of the simulator? I can talk about what's improving a simulator, but we need to measure that the decisions have been made better".

4.4. Maintain a cohesive approach to development and use through governance, capacity building and engagement

The final principle aims to ensure that future development and utilisation of fire simulators is cohesive, strategic and aligns with the varied needs of users in an efficient and equitable way. This may be achieved

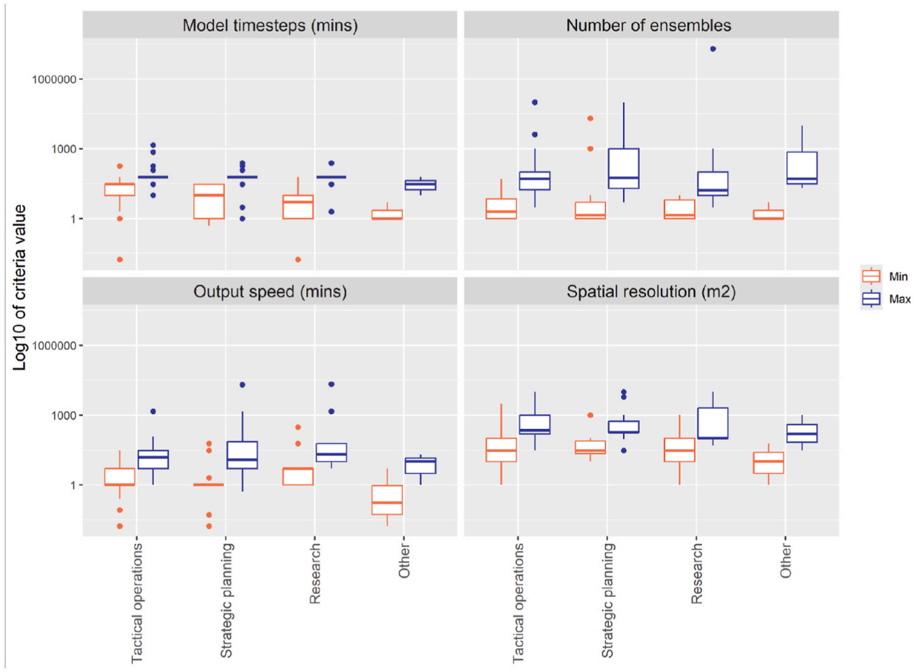

Fig. 4. Performance benchmarks for four aspects of simulator tools: model timesteps, number of ensembles, output speed, and spatial resolution. The y axis is log10 transformed to allow comparison. Orange boxplots show minimum requirements; blue boxplots show maximum. Results are separated by use case. The lower and upper hinges correspond to the first and third quartiles. Whiskers extend up to 1.5 x the interquartile range.

through continued discussion and clarification of the future vision for fire simulators within user communities, particularly as the technology evolves. Addressing shifting simulation goals and priorities for each of the four principles outlined in this paper could simultaneously enhance user-engagement and knowledge sharing (Fisher et al., 2021; Cortes et al., 2024). The benefits of engagement are well documented (Pacheco et al., 2015; Reiter et al., 2018; Henden et al., 2020; Calkin et al., 2021). Continuous consultation ensures that users are supported rather than 'driven' by new technology (Carver and Turoff, 2007), increasing its relevance and therefore adoption (Merritt et al., 2017).

Inclusive consultation also requires concerted efforts to involve stakeholders who may have less capacity to participate in engagement processes and typically hold less influence over decision-making. For example, demographic characteristics such as a participant's jurisdiction and location (often an indicator of their geographical remoteness) were frequently identified as having implications for user requirements. These requirements related to factors such as computing power, data capacities, internet reception, as well as staff skill, experience, and turnover. Similarly, fire regimes and risk profiles were highly variable and context dependent. Smaller jurisdictions were also often said to have little to no capacity to engage in conversations that affect the technology that they use. This disparity in engagement points to a major hurdle for effective and inclusive simulator development in Australia.

The variability in requirements across jurisdictions and locations was often linked to broader debates about governance of simulators and simulator nationalisation. Whilst opportunities for cross-jurisdictional collaboration and cohesive strategy were frequently considered by participants, there was little consensus around whether formal or informal governance structures are required and who would be responsible for funding and implementing such structures. Additionally, due to historical biases and cultural differences between jurisdictions and institutions, ideas about the best way forward were predicted to be similarly biased. As one interview participant indicated, "*… regardless of how good or bad [a simulator is], it's just not going to be accepted depending on where it's come from*" (NSW-SRA1). Forming cross-jurisdictional agreement may therefore prove challenging.

As has been mentioned throughout this study, many of the technical and contextual factors put forward throughout engagement cannot be addressed with a single, unified approach. The lack of consensus was, in part, due to the interconnectedness of factors. It also stems from the reality that fire simulators form part of a social activity, with user preferences, values, and expectations shaping a tool's design, uptake and acceptance (Neale, 2016). Participant views about certain factors were often brought back to critical debates such as how trade-offs in simulator development should be made and who gets to make them; who should be allowed access to simulators and their outputs; how resources and capacities should be distributed across jurisdictions and throughout the sector; and who holds responsibility and accountability for simulator outputs. The latter poses a particular concern given rapid changes to technology and evolving human-computer interactions.

While it is valuable to identify the specific technical and contextual factors important to users, successful development and use of fire simulators cannot focus solely on improving those factors in isolation. Engaging with these larger debates is an inevitable part of decision-making going forward. They may affect not only design, development and user support, but also hold implications for social-licence and legal accountability. Numerous participants, for example, noted that simulator users and audiences may need to justify their actions in a court or coronial inquiry. Sector-wide reflection and consensus building on the appropriate role of science and technology, and the influence of these complex contextual factors, can give managers the conceptual tools to engage with these considerations as they arise (Jasanoff, 2005; Neale & May 2018). Nussbaumer et al. (2023), for instance, established a framework for decision support systems for emergency management, suggesting how to build engagement with value-based considerations into a system and how to train staff appropriately to do so. Ongoing consideration of the values that guide fire simulator development and use in Australia may help to integrate a level of critical reflection into processes that govern the design, development and uptake of these technologies.

5. Limitations and future opportunities

This research did not evaluate specific fire simulators, and our

Table 5
Principles for wildfire simulator development and utilisation.

Principle	Principle components	Actions to Consider
Drive simulator performance through improved modelling and data	*Expand and improve simulator capability*	• Maintain a pipeline of short- and long-term improvements to models and modules
	Improve the availability and quality of input data	• Expand and improve existing input data collection methods, including via automation where applicable • Improve access and useability of input data • Provide guidance on influence of input data scale and resolution on model performance
Improve usability of fire simulator software and hardware	*Ensure user-centric design in development of simulator interfaces*	• Develop front-end user interfaces to reflect new technology and scientific knowledge • Develop intuitive and efficient workflows • Maintain consistency, where possible, across versions and updates • Maximise user customisability (e.g. dual modes for 'general' and 'expert' users) • Build troubleshooting support and feedback into simulators (e.g. error prompts)
	Improve stability and usability of hardware and infrastructure	• Maintain a pipeline of short- and long-term improvements to infrastructure • Provide offline and low resource (e.g. data, memory, computing power) alternatives for simulators
	Improve interpretability and communicability of outputs	• Improve transparency and customisability of outputs • Develop a standardised reporting format for outputs that includes model assumptions and uncertainty • Provide training and support in interpreting simulator output for audiences
	Provide comprehensive support for simulator users and audiences	• Ensure diverse support options are available for users and audiences • Agree upon whether training should be nationalised or accredited
Adopt a comprehensive and transparent approach to validation and verification	*Establish performance standards*	• Establish standards for evaluation, verification and validation of simulators, models and data • Establish a mechanism for reviewing and updating standards alongside evolving technology and contexts of simulator use
	Emphasise transparency and traceability	• Establish guidelines or expectations for documenting simulator function, simulator use processes and simulator outcomes • Automatically collect data on the simulator use process when running a simulation
	Establish outcome-oriented evaluation	• Understand the effect of simulator outputs on fire management outcomes • Develop a process for reporting simulator outcomes post-event
Maintain a cohesive approach to development and use through governance, capacity building and engagement	*Establish effective governance*	• Establish a clear long-term vision for future research, development and use of simulators • Develop strategies for the integration of new science, capacity building, communication, and infrastructure • Reflect on whether current tools are appropriate for intended purposes or if alternatives may be better suited • Consider the benefit to cost ratio of all proposed development and change • Engage with ethical questions, for example, around access, resourcing, responsibility, and human-computer interactions
	Prioritise engagement	• Develop formal and informal mechanisms to engage a wide range of stakeholders at all stages of development to meet their evolving needs • Reflect upon the logistics and resources needed for ongoing, long-term stakeholder engagement
	Build capacity	• Develop strategies for capacity building tailored to user needs

findings should not be used to prescribe recommendations for any specific simulator. Further consideration of simulator development may benefit from targeted assessments of the needs of specific jurisdictions, use cases or simulators. Additional factors worth exploring include funding, resources, organisational structure, data quality, policy and culture, and the needs of audiences and other groups. We aimed for adequate representation of use cases and jurisdictions, but our questionnaire received many more responses from Tactical (37) than Strategic (18), Research (9) and Other (3) users. This may reflect the proportions of use cases but suggests caution in interpreting findings regarding less well-represented use cases. The same point applies to jurisdictional representation. Equitable engagement requires strategies for overcoming barriers to engagement that certain stakeholders face. These results are specific to Australia and care should be taken in transferring findings to other countries.

CRediT authorship contribution statement

Caitlin Symon: Writing – review & editing, Writing – original draft, Investigation, Conceptualization. **Timothy Neale:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Conceptualization. **Gabrielle Miller:** Writing – review & editing, Writing – original draft, Investigation. **Alexander I. Filkov:** Writing – review & editing, Writing – original draft, Investigation, Conceptualization. **Kate**

A. Parkins: Writing – review & editing, Writing – original draft, Investigation, Conceptualization. **Erica Marshall:** Writing – review & editing, Writing – original draft, Investigation, Conceptualization. **Trent D. Penman:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Conceptualization. **Hamish Clarke:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of funding

This research was funded and supported by Natural Hazards Research Australia. HC receives funding from the Westpac Scholars Trust via a Westpac Research Fellowship.

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 acknowledge the Traditional Custodians of the lands where we live and work. We recognise that these lands and waters have always been places of research. We deeply acknowledge the ongoing work by Indigenous people across Australia to nurture resilient communities and landscapes. We pay our respect to Elders, past, present and emerging.

Thanks to the Project Management Team of Thomas Duff (Victorian Country Fire Authority), Simon Heemstra (NSW Rural Fire Service), David Field (NSW Rural Fire Service), John Bally (AFAC), and George Goddard and colleagues at Natural Hazards Research Australia for supporting this work.

We would like to thank the Australasian Fire and Emergency Services Authorities Council's (AFAC) predictive Services Group and members of the Australian fire simulator user community for their support of this research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pyro.2025.100001.

Data availability

The authors do not have permission to share data.

References

Albini FA. Estimating Wildfire Behavior and Effects, vol. 30. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station; 1976.

Alexander ME, Cruz MG. Are the applications of wildland fire behaviour models getting ahead of their evaluation again? Environ. Model. Software 2013a;41:65–71. https://doi.org/10.1016/j.envsoft.2012.11.001.

Alexander ME, Cruz MG. Limitations on the accuracy of model predictions of wildland fire behaviour: a state-of-the-knowledge overview. For. Chron. 2013b;89(3):372–83. https://doi.org/10.5558/tfc2013-067.

Bainbridge L. Ironies of automation. In: Analysis, Design and Evaluation of man–machine Systems. Pergamon; 1983. p. 129–35. https://doi.org/10.1016/B978-0-08-029348-6.50026-9.

Begg C, Dwyer G, Neale T, Pollock I. Established and emerging uses of predictive services in Victoria. Retrieved from Melbourne: https://www.naturalhazards.com.au/sites/default/files/2022-03/established_and_emerging_uses_of_predictive_services_black_summer_final_report.pdf; 2021.

Beven K, Lane S. On (in) validating environmental models. 1. Principles for formulating a turing-like test for determining when a model is fit-for purpose. Hydrol. Process. 2022;36(10):e14704. https://doi.org/10.1002/hyp.14704.

Borup M, Brown N, Konrad K, Van Lente H. The sociology of expectations in science and technology. Technol. Anal. Strat. Manag. 2006;18(3–4):285–98. https://doi.org/10.1080/09537320600777002.

Bowman DMJS, Kolden CA, Abatzoglou JT, et al. Vegetation fires in the anthropocene. Nat. Rev. Earth Environ. 2020;1:500–15. https://doi.org/10.1038/s43017-020-0085-3.

Calkin DE, O'Connor CD, Thompson MP, Stratton RD. Strategic wildfire response decision support and the risk management assistance program. Forests 2021;12(10):1407. https://doi.org/10.3390/f12101407.

Carver L, Turoff M. Human-computer interaction: the human and computer as a team in emergency management information systems. Commun. ACM 2007;50(3):33–8.

Chen N, Liu W, Bai R, Chen A. Application of computational intelligence technologies in emergency management: a literature review. Artif. Intell. Rev. 2019;52(3):2131–68. https://doi.org/10.1007/s10462-017-9589-8.

Cheong L, Bleisch S, Kealy A, Tolhurst K, Wilkening T, Duckham M. Evaluating the impact of visualization of wildfire hazard upon decision-making under uncertainty. Int. J. Geogr. Inf. Sci. 2016;30(7):1377–404. https://doi.org/10.1080/13658816.2015.1131829.

Colavito M. The human dimensions of spatial, pre-wildfire planning decision support systems: a review of barriers, facilitators, and recommendations. Forests 2021 2021;12:483. https://doi.org/10.3390/f12040483.

Cortes CAT, Thurow C, Ong A, Sharples JS, Bednarz T, Stevens G, Del Favero D. Analysis of wildfire visualization systems for research and training: are they up for the challenge of the current state of wildfires? IEEE Transactions on Visualisation and Computer Graphics 2024;30(7):4285–303. https://doi.org/10.1109/tvcg.2023.3258440.

Cruz MG, Sullivan AL, Leonard R, Malkin S, Matthews S, Gould JS, McCaw WL, Alexander ME. Fire behaviour knowledge in Australia: a synthesis of disciplinary and stakeholder knowledge on fire spread prediction capability and application. In: Client Report No. EP145189. Canberra, ACT: CSIRO Ecosystems Sciences and CSIRO Digital Productivity and Services Flagship; 2014.

Duff TJ, Cawson JG, Cirulis B, Nyman P, Sheridan GJ, Tolhurst KG. Conditional performance evaluation: using wildfire observations for systematic fire simulator development. Forests 2018;9(4):189. https://doi.org/10.3390/f9040189.

Fearnley CJ, McGuire WJ, Davies G, Twigg J. Standardisation of the USGS Volcano Alert Level System (VALS): analysis and ramifications. Bull. Volcanol. 2012;74(9):2023–36. https://doi.org/10.1007/s00445-012-0645-6.

Filkov AI, Duff TJ, Penman TD. Improving fire behaviour data obtained from wildfires. Forests 2018;9(2):81. https://doi.org/10.3390/f9020081.

Filkov AI, Ngo T, Matthews S, Telfer S, Penman TD. Impact of Australia's catastrophic 2019/20 bushfire season on communities and environment. Retrospective analysis and current trends. Journal of safety science and resilience 2020;1(1):44–56. https://doi.org/10.1016/j.jnlssr.2020.06.009.

Fisher R, Heckbert S, Garnett S. Reframing wildfire simulations for understanding complex human–landscape interactions in cross-cultural contexts: a case study from Northern Australia. Fire 2021;4(3):46. https://doi.org/10.3390/fire4030046.

Fox-Hughes P, Bridge C, Faggian N, Jolly C, Matthews S, Ebert E, et al. An evaluation of wildland fire simulators used operationally in Australia. Int. J. Wildland Fire 2024;33(4). https://doi.org/10.1071/WF23028.

Garcia C, Fearnley CJ. Evaluating critical links in early warning systems for natural hazards. Environ. Hazards 2012;11(2):123–37. https://doi.org/10.1080/17477891.2011.609877.

Ghali R, Akhloufi MA. Deep learning approaches for wildland fires using satellite remote sensing data: detection, mapping, and prediction. Fire 2023;6:192. https://doi.org/10.3390/fire6050192.

Goubran L, Parush A, Whitehead A. Modelling information flow and situational awareness in wild fire response operations. In: Yamamoto S, editor. Human Interface and the Management of Information: Applications and Services. Lecture Notes in Computer Science(1), 9735. Cham: Springer; 2016. https://doi.org/10.1007/978-3-319-40397-7_2. HIMI 2016.

Hamilton SH, Pollino CA, Stratford DS, Fu B, Jakeman AJ. Fit-for-purpose environmental modeling: targeting the intersection of usability, reliability and feasibility. Environ. Model. Software 2022;148:105278. https://doi.org/10.1016/j.envsoft.2021.105278.

Henden JA, Ims RA, Yoccoz NG, Asbjørnsen EJ, Stien A, Mellard JP, et al. End-user involvement to improve predictions and management of populations with complex dynamics and multiple drivers. Ecol. Appl. 2020;30(6):e02120. https://doi.org/10.1002/eap.2120.

Jain P, Coogan SC, Subramanian SG, Crowley M, Taylor S, Flannigan MD. A review of machine learning applications in wildfire science and management. Environ. Rev. 2020;28(4):478–505. https://doi.org/10.1139/er-2020-0019.

Jasanoff S. Technologies of humility: citizen participation in governing science. In: Bogner A, Torgersen H, editors. Wozu Experten? VS Verlag für Sozialwissenschaften; 2005. https://doi.org/10.1007/978-3-322-80692-5_17.

Jones MW, Abatzoglou JT, Veraverbeke S, Andela N, Lasslop G, Forkel M, et al. Global and regional trends and drivers of fire under climate change. Rev. Geophys. 2022;60(3):e2020RG000726. https://doi.org/10.1029/2020RG000726.

Kelso JK, Mellor D, Murphy ME, Milne GJ. Techniques for evaluating wildfire simulators via the simulation of historical fires using the Australis simulator. Int. J. Wildland Fire 2015;24(6):784–97. https://doi.org/10.1071/WF14047.

McIntosh BS, Ascough II JC, Twery M, Chew J, Elmahdi A, Haase D, et al. Environmental decision support systems (EDSS) development–challenges and best practices. Environ. Model. Software 2011;26(12):1389–402. https://doi.org/10.1016/j.envsoft.2011.09.009.

McFayden CB, Johnston LM, MacPherson L, Sloane M, Hope E, Crowley M, et al. A perspective and survey on the implementation and uptake of tools to support decision-making in Canadian wildland fire management. For. Chron. 2024;100(2):165–79. https://doi.org/10.5558/tfc2024-017.

Merritt WS, Fu B, Ticehurst JL, El Sawah S, Vigiak O, Roberts AM, et al. Realizing modelling outcomes: a synthesis of success factors and their use in a retrospective analysis of 15 Australian water resource projects. Environ. Model. Software 2017;94: 63–72. https://doi.org/10.1016/j.envsoft.2017.03.021.

Miller C, Hilton J, Sullivan A, Prakash M. SPARK – a bushfire spread prediction tool. In: Denzer R, Argent RM, Schimak G, Hřebíček J, editors. Environmental Software Systems. Infrastructures, Services and Applications. ISESS 2015. IFIP Advances in Information and Communication Technology, vol. 448. Cham: Springer; 2015. https://doi.org/10.1007/978-3-319-15994-2_26.

Morrison R, Kuligowski E, Dootson P, Griffin AL, Perry P, Pupedis G, et al. Understanding the challenges in bushfire map use and effective decision-making amongst the Australian public. Int. J. Wildland Fire 2024;33(10). https://doi.org/10.1071/WF24071.

Neale T. Burning anticipation: Wildfire, risk mitigation and simulation modelling in Victoria, Australia. Environ. Plann. 2016;48(10):2026–45.

Neale T, May D. Bushfire simulators and analysis in Australia: insights into an emerging sociotechnical practice. Environ. Hazards 2018;17(3):200–18. https://doi.org/10.1080/17477891.2017.1410462.

Neale T, Vergani M, Begg C, Kilinc M, Wouters M, Harris S. 'Any prediction is better than none'? A study of the perceptions of fire behaviour analysis users in Australia. Int. J. Wildland Fire 2021;30(12):946–53.

Nussbaumer A, Pope A, Neville K. A framework for applying ethics-by-design to decision support systems for emergency management. Inf. Syst. J. 2023;33(1):34–55. https://doi.org/10.1111/isj.12350.

Onnasch L, Wickens CD, Li H, Manzey D. Human performance consequences of stages and levels of automation: an integrated meta-analysis. Human factors 2014;56(3): 476–88. https://doi.org/10.1177/0018720813501549.

Opperman T, Gould J, Finney M, Tymstra C. Applying fire spread simulators in New Zealand and Australia: results from an international seminar. In: Andrews PL, Butler BW, editors. Fuels Management – How to Measure Success: Conference Proceedings', 28–30 March 2006, vol. 41. Fort Collins, CO Portland, OR: USDA Forest Service, Rocky Mountain Research Station; 2006. p. 201–12.

Pacheco AP, Claro J, Fernandes PM, de Neufville R, Oliveira TM, Borges JG, Rodrigues JC. Cohesive fire management within an uncertain environment: a review of risk handling and decision support systems. For. Ecol. Manag. 2015;347:1–17. https://doi.org/10.1016/j.foreco.2015.02.033.

Parker WS. Model evaluation: an Adequacy-for-Purpose view. Philos. Sci. 2020;87(3): 457–77. https://doi.org/10.1086/708691.

Parisien MA, Dawe DA, Miller C, Stockdale CA, Armitage OB. Applications of simulation-based burn probability modelling: a review. Int. J. Wildland Fire 2019;28(12): 913–26. https://doi.org/10.1071/WF19069.

Plucinski MP, Sullivan AL, Rucinski CJ, Prakash M. Improving the reliability and utility of operational bushfire behaviour predictions in Australian vegetation. Environ. Model. Software 2017;91:1–12. https://doi.org/10.1016/j.envsoft.2017.01.019.

Rapp C, Rabung E, Wilson R, Toman E. Wildfire decision support tools: an exploratory study of use in the United States. Int. J. Wildland Fire 2020;29(7):581–94. https://doi.org/10.1071/WF19131.

Reiter D, Meyer W, Parrott L, Baker D, Grace P. Increasing the effectiveness of environmental decision support systems: lessons from climate change adaptation projects in Canada and Australia. Reg. Environ. Change 2018;18:1173–84. https://doi.org/10.1007/s10113-017-1255-9.

Scolobig A, Potter S, Kox T, Kaltenberger R, Weyrich P, Chasco J, Rana B. Connecting warning with decision and action: a partnership of communicators and users. In: Golding B, editor. Towards the "Perfect" Weather Warning: Bridging Disciplinary Gaps Through Partnership and Communication. Cham: Springer International Publishing; 2022. p. 47–85.

Slijepcevic A, Tolhurst K, Fogarty L. Fire behaviour analyst roles and responsibility in bushfire management - how to make the best use of these skills. In: Proceedings of AFAC Conference 2008; 2008. Adelaide. 7pp.

Strauch B. Ironies of automation: still unresolved after all these years. IEEE Transactions on Human-Machine Systems 2017. https://doi.org/10.1109/THMS.2017.2732506. 48(5), 419-433.

Sullivan AL. Wildland surface fire spread modelling, 1990-2007. 3: simulation and mathematical analogue models. Int. J. Wildland Fire 2009. https://doi.org/10.48550/arXiv.0706.4130. 18(4), 387-403.

Van Wagner CE. Two Solitudes in Forest Fire Research. Canadian Forestry Service, Petawawa Forest Experiment Station, Chalk River, Ontario. 1971. p. 7. Information Report PS-X-29.

Van Wagner CE. Fire behavior modelling – how to blend art and science. In: Donoghue LR, Martin RE, editors. Proceedings of the Eighth Conference on Fire and Forest Meteorology. Soc. Amer. For., Bethesda, MD: SAF Publ; 1985. p. 3–5. 85-04.

Walling E, Vaneeckhaute C. Developing successful environmental decision support systems: challenges and best practices. J. Environ. Manag. 2020;264:110513. https://doi.org/10.1016/j.jenvman.2020.110513.