What makes a good fire simulator?

Final Report

Caitlin Symon[1], Timothy Neale[2], Gabrielle Miller[2], Alex Filkov[1], Kate Parkins[1], Erica Marshall[1], Trent Penman[1] and Hamish Clarke[1]

1. University of Melbourne, Victoria, 2. Deakin University, Victoria

Version	Release history	Date
1.0	Initial release of document	05/08/2025

Australian Government

Natural Hazards Research Australia receives grant funding from the Australian Government.

We acknowledge the Traditional Custodians across all the lands on which we live and work, and we pay our respects to Elders both past, present and emerging. We recognise that these lands and waters have always been places of teaching, research and learning.

Publisher:
Natural Hazards Research Australia

ISBN: 978-1-923057-39-5

Report number: 59.2025

August 2025

Cover: Flare Wildfire Research

Table of contents

Acknowledgement of Country

This report was developed on the traditional lands of the Wurundjeri people of the Kulin nation. The authors pay our respect to Wurundjeri elders past and present as well as Traditional Owners throughout Australia. The authors recognise that sovereignty was never ceded. The authors acknowledge their continuing connection to land, waters and community and we acknowledge the long history of reading the land and learning from Country and of working with and alongside fire. The authors also acknowledge all Indigenous people who shared their knowledge and insights for the benefit of the project.

Thank you to research participants

Special thanks to end-user representatives, Thomas Duff (Country Fire Authority (CFA)), David Field (Rural Fire Service (RFS)), Simon Heemstra (RFS) and John Bally (Australasian Fire and Emergency Service Authorities Council (AFAC)) who formed part of the project management committee, along with Natural Hazards Research Australia representatives George Goddard, Kat Haynes and Supriya Gurung. Thank you also to the AFAC Predictive Services Group's Spark-specific working groups for their early input.

Thank you to fire simulator users and developers across Australia who participated in interviews, workshops, questionnaire or otherwise contributed feedback for the project.

Thank you to the FLARE Wildfire Research team, School of Agriculture, Food and Ecosystem Sciences, University of Melbourne and the School of Humanities and Social Sciences, Deakin University.

This research received approval from Deakin University’s Faculty of Arts & Education Human Ethics Advisory Group (approval: HAE-23-031) and was conducted in accord with the approved methods. Research data is stored securely at Deakin University in accordance with Human Ethics Advisory Group requirements and will be held for five years with access restricted to researchers named on the application.

Executive summary

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 are fit-for-purpose.

To support the future development and use of fire simulators in Australia, a comprehensive engagement process with simulator users was undertaken. The focus was understanding users and their priorities and identifying current gaps and issues in simulator design and development. Similarities and differences among the needs and expectations of different simulator users were identified and investigated whether those similarities can inform benchmarks for the development and evaluation of future simulators. Results include:

- Simulator user needs are diverse, context dependent, evolving and resistant to simplification.
- Despite this complexity, two strong themes emerged regardless of individual differences:
    - Social factors are critical i.e. addressing technical simulator criteria is necessary but insufficient.
    - Trust in and support for simulators and their outputs depend upon a combination of robust research; consistent and ongoing engagement and communication between stakeholders; and proven reliability and success of the tool.
- Several criteria were identified as consistently relevant to ensuring fit-for-purpose fire simulators. These were ease of use, speed, configurability, versatility, robustness of modelling framework, effectiveness of software framework, handling of inputs, handling of outputs, scale, validation, support, trust, compatibility and value for money.
- There were some clear differences between *Tactical*, *Strategic* and *Research* uses of fire simulators. For example, Tactical users prioritised *speed*, *ease of use* and *operational support*, while Strategic and Research users valued *configurability* and a *robust modelling framework*, with Strategic users also valuing *high quality outputs*. Overall, however, use case was often a weak predictor of the relative importance of key criteria.
- Highly diverse fire regimes and risk profiles and significant differences in organisational and regional resources and capacity, make *Jurisdiction* and *Geography* important proxies for fire simulator user needs.
- There was some support for generalising user requirements and creating objective benchmarks to inform development. However, the complexity, diversity and overlap of requirements among users and use cases make this challenging. There is also a need to consider factors beyond user needs, including financing, resources, data, organisational structure, policy and culture and the needs of audiences and other groups.
- A set of principles may be a more flexible way to ensure simulators meet user needs and are fit-for-purpose. These principles could be used to guide simulator development and use, while being adapted to meet changing local, organisational and sectoral needs. Within Australia's simulator 'ecosystem', some of these principles are already being implemented while others may require more work. Principles include:
    - Drive simulator performance capability through improved fire behaviour modelling and improved quality and coverage of input data.
    - Improve usability of fire simulator software and hardware, including platform stability, outputs that facilitate effective communication with audiences and ongoing support for users and partners.
    - Adopt a comprehensive and transparent approach to validation and verification, applicable to simulators and their inputs and extending beyond accuracy to outcome-based evaluation.
    - Maintain a cohesive approach to development and use through effective governance, capacity building and inclusive and ongoing consultation between users, audiences and developers.

End-user statement

Thomas Duff - Fire Risk, Research and Community Preparedness, CFA Victoria.

An understanding of user-needs critical for ensuring that decision support tools are fit for purpose. This project was designed to engage with users of fire simulators to gain a better understanding of what they use simulators for and – in their opinion – 'what makes a good simulator'. This information will be important to support the development of cutting-edge fire simulator systems for uses including real-time fire decision making, identification of landscape risk and gaining insight into the processes and mechanisms driving fire impacts.

Introduction

About fire simulators

Fire is an integral part of Australia's ecological and cultural landscape, with a complex interplay between climate, geology, energy and nutrient cycles and active cultures of fire use and management by Indigenous Australians (Kelly et al. 2023). Many species and ecosystems across Australia have evolved alongside fire, forming resistances from or dependencies on it. Particularly since colonisation however, changes to land-use and management, population size and distribution, exclusion of fire from certain landscapes and changing climate have altered fire regimes in Australia (Bowman et al. 2020). In recent years, bushfire[1] events in south-eastern Australia have had significant social, ecological and financial costs (Filkov et al. 2020). Over the coming decades, these costs, along with the burden placed on emergency management systems, infrastructure and social services, are set to grow, as many parts of Australia and the planet face increasingly extreme bushfire risk under climate change (Bowman et al. 2017; Cunningham et al. 2024; Di Virgilio et al., 2019; Doerr & Santín 2016). Fire management agencies play a central role in our response to these risks.

Within fire management and research, fire simulators have become an important tool to understand, predict and communicate information about fire over short-term (Tactical) and long-term (Strategic and Research) time horizons. Fire simulators are computer programs used to depict key processes involved in the behaviour and spread of fire through a landscape. At their heart are physical and empirical models of fire behaviour and their inputs include information about terrain, weather, vegetation and assets. Fire simulators allow users to test scenarios that otherwise cannot easily be tested in real life due to constraints around cost and resources, safety and time. Fire simulators only gained widespread uptake over the past two decades and the current reliance on detailed information from simulators is relatively new. This is, in part, due to the improvement of simulation technology, but also because of the new challenges and risks posed by bushfire.

Alongside the technical factors involved in fire simulation, there are a range of complex organisational and social factors that also impact fire simulator development and use. The capacities of fire researchers and management agencies to develop and use these tools vary widely, but in general, the bushfire sector is often resource limited and time poor and staff must make trade-offs and compromises in how they use simulators to best support their work. These compromises firstly occur in day-to-day use of simulators. For example, when developing an output, users must balance the time taken with the amount of detail, accuracy and communicability of results. Trade-offs are also an integral part of the implementation of simulation systems. Factors such as cost, capability, computing and data requirements and staff capacities must be weighed up. Active debate and research have surrounded the improvement of fire simulators, including the development of new models and simulator programs. There are currently, however, no agreed upon standards for simulator development and there are knowledge gaps about how best to consolidate the wide-ranging views on how funding and development should be prioritised.

Fire simulator use and evaluation

Cruz et al. (2018) note that the mathematical models of the forward rate of spread of a fire have been under development for close to a century. These models underpinned the original simulators, which were not computer programs but rather paper-based forms and slide rules designed for field use (see Plucinski et al. 2017 for more information). Evaluation of computer fire simulators in Australia dates back at least forty years, to when Green et al. (1983) evaluated a simple simulator's ability to represent the shape of experimental fire burned areas. Fire simulators have grown in complexity and application since then. Some authors refer to the fire simulator as the central piece of a larger system (Kelso et al. 2015), highlighting the importance of input data, specific mathematical modules and simulation software.

[1] We use the term bushfire to refer to unplanned vegetation fires. Other terms include wildfire, wildland fire, grass fire and forest fire.

There have been a number of studies that aimed for a relatively comprehensive characterisation of the performance and state of fire simulators, in contrast to the more targeted studies noted above. Validation methodology studies highlight the diverse range of fire simulator outputs that can be compared with observations, proposing metrics as well as frameworks for developing and selecting these metrics (Filippi et al. 2014; Kelso et al. 2015; Duff et al. 2016; Humber et al. 2020). Evaluation studies systematically compare the performance of different fire simulators. A landmark evaluation of Australian operational fire simulators noted the importance of ensemble runs, modular and transparent models, standard outputs for analysis and evaluation and national case study hosting (Faggian et al. 2017; Fox-Hughes et al. 2023). A recurring theme of this and other evaluation studies is the need for better monitoring of fires and particularly of extreme fire behaviour, to support the development of reconstructions and improved evaluation (Duff et al. 2018; Cruz et al. 2018). Improved fire monitoring is regularly identified as a high priority in reviews of various simulator types (Cruz et al. 2014; Bakshaii & Johnson 2018; Wegrzynski & Lipecki 2018; Morvan et al. 2022; Sun et al. 2023; Zacharakis & Tzihrintzis 2023; Ghodrat et al. 2023). Other priorities highlighted by these reviews include multi-scale modelling, improved canopy & terrain roughness models; assimilation of weather data for computational fluid dynamics models and improvements in fuel mapping and other inputs including wind.

In addition to these systematic studies of fire simulator performance, many studies that use fire simulators implicitly contribute to our understanding of simulator performance, limitations and areas of improvement (Salis et al. 2016; Penman et al. 2020; Penman et al. 2022), vertical atmospheric structure (Ozaki et al. 2022), fire-atmosphere coupling (Peace et al. 2021) and the use of active fire data from satellite observations for model validation (Sa et al. 2017). Fire simulator evaluation extends to new applications of models (Ujjwal et al. 2022; Gonzalez-Olabarria et al. 23) including model visualisations (Cortes et al. 2023). Finally, although there are a limited number of fire behaviour models that underpin operational fire simulators in Australia, the continued development of new simulators and software to deploy these models suggests that flexibility in model choice will be important for fire simulation systems of the future (Katan and Perez 2023; Pais et al 2023; Prichard et al. 2023; Schumaker et al 2022, Wu et al. 2023).

In their survey and interviews with 32 fire behaviour analysts, Cruz et al. (2014) identified priorities for future inclusion in fire simulators, such as night-time fire behaviour, build-up times and transitions to convective-driven fires. A more recent list of knowledge gaps can be found in Plucinski et al. (2017), including fire spread away from the head-fire, initial fire growth phase duration, fuel types that lack fire behaviour or flame characteristic models, short-range spotting, prescribed burning, fire spread on slopes and complex terrain and fire spread under extreme conditions. Cruz et al. (2014) stands out for its engagement with fire simulator developers, recording many fire simulator features including fire behaviour models incorporated, fuel types incorporated, fire perimeter propagation method, simulation time for a 100,000 ha fire. They also recorded a number of software-related attributes: computer language, cross-platform availability, source code control, integration and testing, system testing, bug tracking.

To date, most evaluation of fire simulators has focused on the ability to sufficiently represent or 'mirror' key aspects of fire behaviour. This is not unreasonable given that the foundation of applied environmental modelling is the representation of environmental phenomena, particularly those aspects that are most relevant to decision makers. This capacity to provide a decent representation of reality sets a minimum performance goal for fire simulators to be useful. Users, however, increasingly recognise a simulator's fitness-for-purpose as an equally important goal alongside its capacity to depict an ever nearer approximation of reality (Beven and Lane 2022; Hamilton et al. 2022). With increasing use of fire simulators in environmental and emergency management, there has been an accompanying increase in research into organisational, cultural and other contextual factors that support or inhibit the effective use of fire simulators. There is an emerging consensus that evaluation must extend beyond the simulator itself and that engagement between simulator developers, users and audiences is critical. Simulation of physical processes related to fire remains, however, a cornerstone of model development, use and evaluation.

Beyond fire behaviour - lessons from social science and applied environmental modelling

There are relatively few studies of the actual development and evaluation of fire simulators to draw upon despite their widespread use. Nonetheless, there are a wealth of studies of the wider category of environmental decision support systems (DSS), such as forestry growth models and flood simulators. This literature on evaluation frameworks and design of decision-support tools consistently states the need to shift focus away from solely judging simulators by their capacity of outputs to represent reality (Parker 2020). Output quality is important, however, there are many additional factors that can also influence whether a simulator has a positive or negative impact on management outcomes (Matthews et al. 2011; Morss et al. 2005; Rayner et al. 2005). Much of the research conducted on decision support systems instead suggests that simulators should be assessed on their fitness-for-purpose (Beven and Lane 2022; Hamilton et al. 2022). Fitness-for-purpose requires a greater understanding of the social and institutional constraints of simulator use because the practical implementation of decision-support systems involves multiple stakeholders and is often highly complex (Pacheco et al. 2015). As the review by Walling and Vaneeckhaute (2020) suggests, uptake of a DSS should be supported by not only integrating users and key audiences into development but also by analysing the social context of their decision-making. Similarly, Hamilton et al. (2019) argue, DSS evaluation needs to be ‘multi-dimensional and multi-perspective’, attentive to the factors influencing the extent to which a DSS is accessible, credible and salient in the eyes of its users and audiences. It is important to develop mechanisms that ensure collaboration and communication between developers, users and audiences as well as capacity building (Merritt et al., 2017; Neale et al., 2021; McIntosh et al. 2011; Rapp et al. 2020).

As in other contexts (Pacheco et al. 2015), expert judgement and intuition are requisite to the implementation of these simulators (Cruz et al. 2014; Plucinski et al. 2017; Neale et al. 2021). This is due to both the persistence of significant ‘knowledge gaps’ in underlying sub models and datasets and the institutionalisation of a class of expert users trained and employed to negotiate these gaps and generate meaningful simulator outputs. In broad alignment with studies of Wildland Fire Decision Support Systems, studies of Australian fire management contexts emphasise the vital importance of trust and credibility to DSS evaluation, underpinned by the credibility of both a DSS and its users to an audience (Neale 2023; Neale and May 2020). In their study of fire behaviour analyst (FBAn) audiences within incident management, for example, Neale et al. (2021) argue that an audiences’ values (e.g., a preference for experiential knowledge) shape the credibility of simulator outputs based on their knowledge of the user. As a result, they call for fire management institutions to invest in fostering social connection between simulator users, their tools and their audiences alongside investments in fire behaviour simulators. This call builds on earlier surveys which highlighted the importance of engagement between developers, FBAns and incident management teams, training and alignment with existing tools and processes (Cruz et al. 2014). Voinov & Gaddis (2008) and Voinov & Bousquet (2010) suggest mechanisms for effectively engaging stakeholders such as participatory modelling. Evaluation should be a distinct, iterative process of model development and use and peer review can also be included as an important research phase of model development. Such procedures and frameworks are necessary for increasing the quality of these tools and ensuring that trust and collaboration are systematically built into simulator development.

Project background and research aims

Our overarching question is: **What makes a good fire simulator according to its user community?**

The initial research proposal was developed by an Australasian Fire and Emergency Service Authorities Council (AFAC) Predictive Services working group, which identified key knowledge gaps in simulator development:

1. To understand how existing fire simulators are currently being used and will be used in the future.
2. To consider the feasibility of developing standard requirements to ensure simulators are fit for purpose.
3. To develop a picture of user priorities for simulator improvement or development.

The initial proposal had the key objectives of determining key use cases for simulators and understanding, for each use case, the following social questions:

- Who are simulator users?
- How are simulators used and what decisions do they support?
- What criteria inform their use?
- How can we measure these criteria?
- What thresholds must each of these criteria meet to be considered 'good enough' for their intended purpose?

From these initial research aims and through consultation with key end-users, the priority objectives of this project were further clarified and refined, leading to two key priorities: to identify the strengths and weaknesses of current systems and provide guidance on how we can improve systems to benefit all users in the future. A third priority was added after consultation with lead end-users, around establishing quantitative benchmarks.

Eventually these priorities were further refined into the following research aims, around which all engagement was built:

a. Understand current practices, issues and priorities of future development for the broad and expanding community of fire simulator users.
b. Establish a set of standards or measurements that reflect simulator fitness-for-purpose requirements to guide future simulator development.

Scope

The project focuses on all **simulators used in or to support fire management** in Australia. The network of stakeholders involved in fire simulators across Australia is vast and therefore to ensure that the knowledge gaps highlighted by AFAC in the initial research proposal were thoroughly and directly responded to, this project focuses specifically on the perspectives of **simulator users**. Simulator users, including developers, are individuals who actively engage with the fire simulator tools and are responsible for developing simulator outputs for audiences[2]. Other factors and stakeholders that are relevant to simulator development are mentioned in this report, however they are not considered in-depth. The recommendations in the report's 'Discussion' section include future research opportunities to build on our findings and engage further with the needs and perspectives of the broader fire simulator community.

[2] In this report users are also referred to as operators and to audiences as end-users or clients.

Purpose of the report

The variability of fire simulator users across Australia means that the recommendations and implications for management presented in this report are not designed as a single, prescriptive set of instructions for simulator developers to follow. Instead, they are general in nature and serve the following purposes:

- To consolidate the vast and variable knowledge and perspectives of fire simulator users across Australia and to document their requirements and priorities for future improvements to fire simulators.
- To provide guidance to the fire management sector for how to strategically and collectively improve fire simulators.
- To serve as a reference document for simulator developers, including as a checklist of important areas for consideration in simulator development that can to sit alongside and inform detailed in-house simulator-specific directives.

Research approach

Project overview

This engagement-focused project proceeded in two phases. Phase 1 focused on the identification of key themes and was designed to identify important areas of consideration for fire simulator development and use, including different purposes for simulator use and the different types of users, existing knowledge and data gaps; simulator performance criteria; limitations of current systems and future challenges. These findings were used to inform stakeholder engagement in Phase 2.

Phase 1 consisted of:

- **A review** of Australian fire simulator literature to ensure that the engagement aspects of the project were informed by the latest research into fire simulators.
- **Semi-structured interviews** to identify issues and themes.
- **Two workshops** to discuss and, where possible, quantify issues

Phase 2 consisted of **an online questionnaire** to quantify the issues and themes raised in Phase 1.

Participant selection

Participants were selected in consultation with the project management team to represent the diversity of the Australian fire simulator user community. This included a mix of jurisdictions and sectors (i.e. management, research and private organisations). Jurisdiction reflects the State or Territory that a participant's work primarily engages with. Participants who worked across jurisdictions, at the national level, were categorised as 'National'. Early consultation suggested three distinct use cases. Within each are many specific uses (Table 1). These were *Tactical* and *Strategic* (generally restricted to fire agencies) and *Research* (within research organisations, fire agencies and consultancies). The category *Other* was used for community education, training and any other uses that did not fit into the first three. Tactical and Strategic simulator use occurs within a decision support (i.e. operational) setting. Research use (a.k.a. knowledge development) of simulators occurs in both research organisations and fire agencies. Whilst users generally associated themselves with one of the use cases, some did not neatly fit these categories. Some participants did not fit into any of the three and were categorised as 'other'.

Tactical

The use of fire simulators for tactical purposes primarily involves assessments of current or potential fire behaviour and spread. Tactical use of simulators results in real-time predictions (or as close as possible) to make rapid, time sensitive decisions regarding resource allocation or resource prioritisation, determining optimal deployment strategies and location of key personnel, as well as helping managers decide how and when risks are communicated to the public and when evacuations need to be considered.

Strategic

Where tactical decision making requires rapid and often faster than-real time outputs, simulator use for strategic purposes is generally less time restrictive and commonly utilised for longer term planning. Strategic planning using fire simulators frequently involves the development of fire management plans/policies, mitigation strategies, fire season resource allocation and prioritisation of key areas for preventative actions such as prescribed burning or other fuel reduction strategies. Fire simulator use for strategic needs is generally undertaken by simulating multiple different fire scenarios and assessing the potential impacts or risks associated with different variables.

Research

Similar to strategic uses, fire simulator use for research purposes is generally less time-restrictive and rarely requires real-time predictions. Use cases for research purposes cover a broad range of applications. The most common use cases include research which focuses on advancing our understanding of fire behaviour, fuel dynamics, environmental factors and other variables which are known to influence fire spread. Through careful exploration of simulators and the key variables, researchers can refine existing predictive models, identify emerging trends and inform future research directions aimed at improving fire management practices. Research into the fundamental aspects of fire dynamics, fuel components and environmental influences facilitates the development of more accurate predictive models, enhancing our understanding of fire behaviour.

Other

Fire simulators are also used: to guide policy development; for training and education purposes (to build the decision-making skills and preparedness of emergency responders through simulated scenarios); and as educational outreach tools for the broader community to raise public awareness about fire risks, circulate current messaging around leaving early or preparing fire action plans.

TABLE 1 - COMMON USES OF FIRE SIMULATORS

TACTICAL	STRATEGIC	RESEARCH	OTHER
Tactical prediction (real-time)	Landscape risk assessment including fuel management and scenario analysis	Research into tactical use cases	Community engagement
Tactical suppression	Readiness at various timescales	Research into strategic use cases	Staff engagement
Back burning planning	New estate siting and layout incl new settlements	Simulator evaluation	Scenario training
Public warnings	Prescribed burning planning	Understanding processes and module development	
Smoke assessment	Scenario training	Fire spread reconstruction	
Evaluation of effectiveness of interventions		Evaluation of effectiveness of intervention and effort	
Readiness			

Interview and workshop participants were selected for their expertise and experience as participant numbers were limited. The questionnaire aimed to reach all simulator users in Australia and included simulator users regardless of experience or skill level. Interview and workshop participants were approached by phone or email. The questionnaire was advertised by email and online (e.g. LinkedIn). The final breakdown of participants is shown in Table 2.

TABLE 2 - SUMMARY OF PROJECT PARTICIPANTS

	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
Queensland	1	3	3	8
South Australia	2	2	4	2
Tasmania	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[3]	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
Otherwise identified	0	0	0	0
Prefer not to say	0	0	0	1
Total	**21**	**26**	**25**	**67**

Literature review

The aim of the literature review was to ensure that the engagement aspects of the project were informed by the latest research into fire simulators. The review was broken into four parts: definitions of different kinds of fire simulator, uses of fire simulators, the development and evaluation of simulators and lessons from other applied environmental models. The section on development and evaluation of simulators considered both technical ('content') and social ('context') aspects. The review focused on peer reviewed scientific journal articles, complemented by grey literature where necessary. The key findings from the literature review have been included throughout the introduction of this report and, in addition, will be submitted as a standalone journal article.

Interviews

Interviews were conducted between 31 July 2023 and 26 October 2023.

The interviews were guided by a schedule of questions designed to engage with several key themes: who is using simulators and for what purposes; key limitations, knowledge and data gaps within existing simulators and their use; simulator performance criteria; and treatment of uncertainty and future challenges in simulator development and use. The interviews were semi-structured to allow space for participants to guide and direct conversation. Discussions centred around the following questions:

- What makes a good fire simulator?
- What are the key criteria for judging between simulators?

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

- How do you know one simulator is better than another?
- Thinking of the fire simulators you are most familiar with, generally speaking, it is more important that they become: more accurate or faster; more accurate at a fine scale or more accurate at a coarse scale; more compatible with other simulators and models or more capable to model all aspects of fire behaviour itself?
- What are the most appropriate paths forward to develop fire simulators into the future?
- What are the key factors in ensuring a new fire simulator is utilised by practitioners like yourself?

Results of the analysis of the interview data were organized according to key emergent themes (thematic codes) that related to the project's key research question, 'What makes a good fire simulator according to its user community?'. These themes informed the design of Workshop 2 and the project questionnaire. Interviewees were categorized into 4 use cases (Tactical Operations [TO; n=7], Strategic planning [SP; n =8], Research [Res; n=6], Community and staff education [CSE; n=1]). Quotes from participants have been attributed using a code to protect their anonymity, revealing 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.).

Workshop 1

Workshop 1 was held online on 9 October 2023. Key topics were fire simulator use cases, performance criteria and decisions about changing fire simulators. Participant responses were recorded using Qualtrics and by note takers. Users were asked to identify specific fire simulator use activities within three broad use cases: Tactical, Strategic and Research. A fourth category contained other use cases such as education and training. Participants were then presented a list of 35 criteria, (Table 3) drawn from informal consultation with fire simulator users and the literature review and were asked to then rate the importance of each criterion on a scale from 1 to 5 (low to high importance). Finally, participants were asked to consider three change scenarios: a new fire simulator input layer, a new version of the currently used simulator and a different fire simulator altogether. For each scenario, participants answered three questions in relation to performance criteria: what would be the minimum % improvement required to recommend a change, what % improvement would make you recommend a change without hesitation and what % decrease in performance could you tolerate before recommending the change not be made?

Workshop 2

Workshop 2 was held online on 26 February 2024. It gathered qualitative data from participants to assess the possibility of and value in quantifying thresholds and benchmarks for fire simulator performance and development. The questions and performance criteria (Table 4) were developed as a consolidated list of the criteria from Workshop 1 (Table 3). The criteria list was shortened based on findings and feedback from Workshop 1 and informal consultation with simulator users.

Participants were presented with three performance scenarios:

1. Unacceptable - Thinking about simulators generally, both existing, in development and yet to emerge, what level of performance would you consider unacceptable?
2. Worth considering change - Thinking about how you use simulators right now and where each criterion is right now, what level of performance would entice you to support making a change? If relevant, distinguish between level of change i.e. new data, new version of same model, entirely new model.
3. Game-changer - What level of performance, if any, would you consider game-changing?

As many of the participants in this workshop were not decisionmakers but instead were the stakeholders responsible for actioning changes to simulator use within an organisation or agency, the questions for Scenarios 2 and 3 were intentionally framed to ask them what changes would mean they support a switch or what factors would facilitate their switching rather than factors that would influence their own decision-making about a simulator.

Participants were asked whether their answer depended on their primary use case (Tactical, Strategic, Research, or Other)` or on other factors and whether their answer could be measured objectively. This line of questioning was designed in response to Workshop 1, where participants had emphasised that quantifying expectations for and improvements to simulators was challenging.

Questionnaire

The questionnaire consisted of 25 questions and was separated into 3 sections:

- Section A: Demographics
- Section B: User requirements
- Section C: User preferences

Section A, Demographics, was used primarily to get a clear picture of who was participating in the questionnaire and to ensure that the results were representative.

The user requirement questions sought to identify a minimum acceptable threshold - the amount, below which, a simulator performance would be unacceptably low for their uses and a maximum necessary threshold - the amount, above which, any increase in performance would have a negligible effect on the quality of their output. The questions in this section questions related to performance factors such as the accuracy of fire characterisation, speed, resolution and ensembles.

The user preference questions asked participants to rank by importance, factors related to user experience, output quality, outcome quality, trustworthiness and the broader 'simulator ecosystem' (important factors external to the simulator tool).

This survey was built in Qualtrics and was sent out to relevant stakeholders across the sector. It was made available online between 1 July 2024 and 2 August 2024.

TABLE 3 - PERFORMANCE CRITERIA FOR FIRE SIMULATORS, FROM WORKSHOP 1

CRITERIA	EXPLANATION INCL. SUB-CRITERIA
EASE OF USE	Including interface
SPEED	Speed of entire process, from data input to usable output
	Speed of specific part of process e.g. fire spread simulation
CONFIGURABILITY	Tweakability, tuneability, tailorability, parameterisability
	Useable by different user types and skill levels
SPECIALISED / UNIVERSAL	Versatile – able to model all aspects of fire behaviour
	Targeted – specialises in specific aspects of fire behaviour
MODELLING FRAMEWORK	Able to run ensembles
	Able to explicitly consider uncertainty
	Deterministic
	Transparency e.g. absence of 'black boxes' and machine learning
SOFTWARE FRAMEWORK	Platform
	Open source
INPUT PROCESS	Inputs – sensitivity to inputs
	Inputs – ease of preparation
OUTPUTS	Outputs – range of detailed spatial outputs
	Outputs – ease of storage and curation
	Outputs – able to be audited
SCALE	Scale – fine spatial resolution
	Scale – spatial extent (e.g. expansion to statewide)
	Scale – fine temporal resolution
	Scale – temporal extent (e.g. multi-day fires, regimes)
VALIDATION	Validation – overall accuracy
	Validation – accuracy for specific vegetation types
	Validation – nature and number of test fires (statistical power)
SUPPORT	Access to training, education, technical support
REPUTATION	Familiarity, awareness
	Respected by peers
	Endorsed by agency / organisation
	Trust in developers
	Trust in underlying research
COMPATIBILITY	Compatibility with agency systems
	Compatibility with agency programs & policy e.g. risk horizons, baselines
	Compatibility with other models e.g. loss & impact models, forecasts
PRICE	Cost

TABLE 4 - PERFORMANCE CRITERIA FOR FIRE SIMULATORS, FROM WORKSHOP 2

Validation	Resolution	Ensembles	Compatibility
Accuracy	Setup	Operational support	Reputation
Speed	Technical outputs	Interpretation support	

Research findings

Interviews

Interviewees reported using fire simulators for diverse fire management purposes across government, industry, community and research fields. Further, several interviewees described fire simulators as a potentially "powerful tool" [QLD-SP1] for purposes that overlap and extend beyond fire management, including corporations such as power and water utilities [Aus-Res1]. Coding and analysis of the project interviews distilled several themes related to one of two domains of activity: simulator use and simulator development (Table 5), the detail of which is outlined below.

TABLE 5 - THEMES OF INTERVIEW ANALYSIS

SIMULATOR USE	SIMULATOR DEVELOPMENT
Purpose and usefulness	Trade-offs and priorities
Expert vs general users	Development challenges and gaps
Interpretation and communication	Collaboration/coordinated approach
Configurability vs automation	Governance
Reputation, trust and confidence	Transparency and documentation
Utilisation challenges and gaps	Evaluation and verification

Simulator use

Purpose and usefulness

Participants across the use cases consistently emphasised that what is needed from a simulator is highly dependent upon the purpose for which it is being used; "it's about a simulator that's useful for a decisionmaker at the time" (Vic-TO1). Interviewees provided different examples of shifting prioritization of criteria depending on the "hat they're wearing" (NSW-SP2), reflecting the different needs and time constraints of the user or decision-maker.

Overwhelmingly, the primary purpose of fire simulators was described as a decision support tool. Several Tactical interviewees stated that the purpose of simulators in an operational context is to do what manual predictions can with greater accuracy and efficiency. There was a strong theme across interviews with Tactical users that "good" simulator practice is essentially tied to the skills, intentions, knowledge and experience of the user.

For planning purposes (e.g. prescribed burning programs, or land use planning), "good" simulators facilitate the production of verifiable and accurate information to justify planning decisions and policy-making. For these purposes, interviewees viewed accuracy (over timeliness) of simulator outputs, especially at finer scales, as critical to its usefulness. The sole interviewee using simulators for community education and training evaluated simulator usefulness in terms of how well simulator outputs can communicate uncertainty to a range of audience types.

Expert vs general users

Responses coded under this theme were divergent; there was variation in perspective and strength of sentiment regarding the kinds of users (experts vs. general) that simulators should be designed for. These views were mostly expressed in response to a question about whether simulators should become more programmable by expert users or more useable by general users. Interviewees perspectives on access appeared to influence the value that they assigned to certain performance criteria including ease of use and configurability.

Some participants advocated for widespread access to simulators, embracing the notion that a diverse range of users within agencies with varying skill levels should have the opportunity to engage with this technology. A minority further envisioned simulators as tools that could empower members of the public to better understand fire risks, thereby enhancing community resilience. Other participants, however, opposed this notion, asserting that simulators should be exclusively reserved for a select group of highly-trained experts. They argued that the risks associated with wider availability of simulators outweighed the potential benefits, concerned that simulators could be misused in ways that led to detrimental impacts on fire-prone communities and decreased confidence and trust in simulators. They emphasised the uncertainties innate to simulators and the need for rigorous training to ensure that those uncertainties are managed.

Interviewees from all use cases broadly agreed improvements to overall ease of use of simulators would be a welcome development. Some participants envisioned the potential future development of an "ideal" simulator with two streams: an "advanced" more configurable model and a "general user" more automated model (discussed in configurable v automated below). These individuals viewed a trade-off between expert and general users as non-essential and suggested the needs of multiple user groups should be factored into future development.
The question of how to train and support the development of users capable of effectively using simulators, disseminating, communicating and interpreting outputs yielded similarly diverse perspectives. There were mixed thoughts on availability of training as a key factor for utilisation. A minority of participants suggested that good UX design precludes a need for some training. Some participants identified barriers to gaining experience or developing skills via learning approaches: agency hierarchies or 'chain of command' can limit opportunities for upskilling of staff, with the lack existing trained staff frequently cited as a utilisation challenge; long time periods between fire seasons or deployments resulting in limited practical experience; and, lack of sufficient consistent and ongoing training to be effective.

Interpretation and communication

Effective interpretation and communication of simulator outputs was highlighted across interviews as a determinant of good utilisation. Several interviewees stated that the purpose, necessary assumptions and limitations involved in running simulations need to be well documented and communicated for outputs to be useful to decision-makers.

There was a widely shared belief that poor communication and interpretation leads to reputational damage and was therefore identified as a key factor determining "bad" simulator use. One participant explained this as:

> *"We'll have a lot of legwork to do to actually try and improve that kind of understanding of simulators. The example for that would be that some people will only ask for manual fire spread predictions because they think the simulator is rubbish. That's not the simulator really, it's just the underlying data in the simulator and it's the person that's run the simulation doesn't communicate that with a whole heap of accompanying context" (WA-TO1).*

Many interviewees described previous examples where a lack of understanding or awareness of simulator capabilities and limitations led to misinterpretation of outputs. The following quote reflects a commonly cited challenge for simulator utilisation (especially among Tactical users) that "too many people think that the simulation output is a reality. And so instead of going to look at what the reality is, they just take the simulation as the reality" (Aus-Res1). Reiterating the purpose of simulators as a decision-making tool, several participants discussed the importance of improving the accuracy of simulator alongside improvements in the communication of simulator limitations and assumptions. This was viewed as critical for building reasonable expectations of outputs and in turn, sufficient level of trust and confidence for outputs to be used in decision-making.

Several interviewees suggested there needs to be a whole-of-agency approach to training to develop awareness and understanding of simulators. Several suggested that there is a need for differentiated training, that includes training of Tactical users but also training for 'downstream' decision-makers in the interpretation and use of outputs. However, while training in communication of simulator outputs was widely seen as critical challenge, many interviewees suggested design and development of "good" simulators should prioritise features that help communicate uncertainties and assumptions within outputs. Other communication challenges identified were the potential for simulators to fall into "the wrong hands" or for the rise of competing information via technology and social media companies. Both of these challenges risk the possibility of agencies losing control of the narrative via third party involvement in dissemination of simulators and simulator outputs. This has the potential to undermining broader trust in simulators [US-Res2]. One participant also highlighted the additional challenge of deciding whether to "[partner] with these [third party] agencies to bring them on board versus… communicating that they're not efficient products that have been released by the [third party] agencies" (Vic-TO1).

Configurability vs automation

The configurability of simulators by users was emphasized as a key factor determining its usefulness, particularly among Tactical users. Simultaneously, some level of automation was widely viewed as a desirable feature, so that "even if the underlying model is complex the tool itself shouldn't add to the sort of mental overhead of the operator" (Aus-Res3).

Both too much configurability and not enough configurability were cited as drivers of poor simulator use. In regions or jurisdictions where there is less need or opportunity for simulator use, ease of use is a high priority. While the full configurability in Spark where "nothing is impossible" is useful for expert or daily users, it is not useful for others, which as one respondent put it, undermines the purpose of simulators to speed up the process (WA-TO1). Configurability was widely cited as a top performance criterion among Tactical users, who described configurability interchangeably with other features such as transparency and verification. Configurability requires transparent modelling frameworks and enables verification and evaluation of simulators outputs. One Tactical user described automated features of simulators as "getting in the way" of the user (TAS-TO1).

Some participants rejected the notion that configurability and automation capabilities must necessarily be traded off in future development of simulators. They suggested that simulators should be designed with two user types: expert user and general user, as Aus-Res1 suggested, "it needs to be possible to automate the simulation process, but it also needs to have an interface where power users can use it."

Reputation, trust and confidence

Interviews explored participants' views regarding the reputation of simulators among users as an important criterion for utilization of simulators. Responses to direct questioning were split regarding 1) the extent to which peer reputation influenced utilisation; and, 2) how simulators developed a reputation.

Reputation as a factor determining good utilisation:

Some interviewees viewed reputation as "less of an issue" for simulator utilisation in Australia due to agency endorsements and a smaller pool of appropriate simulators: "We're in a situation where there's not really competing options for models in Australia" (SA-TO1). These respondents generally believed that the few simulators available or endorsed by agencies are appropriate for use. However, the interviews suggest that different agency cultures, histories of use and engagement with simulator development have manifested different reputational issues. A participant from a smaller jurisdiction noted that the reputation of simulators was a bigger issue than in a larger jurisdiction such as Victoria, where simulators have been used and integrated for longer. Further, some suggested jurisdictional and institutional biases are having a differential influence on the reputation of simulators among users across Australia: "There's clearly jurisdictional bias in Australia and also probably institutional bias as well. And so regardless of how good or bad some are, it's just not going to be accepted depending on where it's come from" (NSW-SP1).

Regardless, the majority of participants viewed peer reputation as a critical factor of "good" simulators in the past and for determining the success of transitioning towards and integrating new simulators in the future. Building trust and confidence in simulator outputs among "downstream" user groups was viewed by many interviewees as critical for simulators to fulfil purpose as a decision-making tool: "obviously reputation is very important because if you want to [make a] decision, you need trust and you build trust by reputation" (Eur-Res1).

> *You want not just a good reputation for the FBAN [Tactical] practitioners, you want it to be widely accepted as something that's useful in a management perspective or have some sort of authority. When you go to someone with an output from it, you want the people that are actually making the decisions to actually believe in what you're giving them" (SA-TO1).*

Further, the development of fire simulators in the future will require publicising simulators' demonstrated value to users. Building reputation and understanding of new fire simulators is a critical pre-cursor for transitioning to new fire simulation models and systems and was discussed in relation to the ongoing roll-out of SPARK: "Reputation and comfort with things is certainly a big issue. Until we get something up to speed… and have a comfort level that it's doing as good a job or better, then I think we won't move on" (Aus-SP1).

Factors affecting reputation, trust and confidence:

Discussion of factors affecting simulators' reputation was widely discussed in terms of 'trust' and 'confidence' in simulator outputs and simulators themselves. Many participants were aligned in suggesting the following criteria build confidence and trust and thereby improve a simulators' reputation among users:

1. Support and training: needed to build user familiarity and confidence with simulators. Proximity to simulator developers through open lines of communication also helps build confidence & trust among users.
2. Transparency and documentation: needed to ensure simulator not seen as a black box. Documentation facilitates users' developing an understanding of assumptions, limitations, inputs and outputs.
3. Actual or perceived shortcomings: some stated that simulators developed poor reputations from actual or perceived shortcomings in accuracy, timeliness, ease of use or ease of inputting data; "If it's got a shitty reputation, then why does it have a shitty reputation? That's probably either because it's hard to use, it's hard to get the right data for it or it's not very accurate or it's slow. So usually there's some aspect that's not ideal with it" (SA-TO1).
4. User error: some rejected the notion that poor reputation stemmed from simulator shortcomings, but rather resulted from user error. For example, several interviewees suggested that an observable reluctance to embrace simulators among some user groups stems from use of outputs without adequate communication and understanding of their limitations and assumptions.

Utilisation challenges and gaps

Many participants noted that challenges and gaps associated with simulator use was not necessarily due to the simulator itself, but rather issues relating to user experience and understanding of simulators, availability and quality of input data and the contexts in which they are used. These issues echo the content of the previous themes.

1. Availability or quality of input data: this was the most widely cited challenge or gap limiting "good" simulator use. The view that "a good model is only as good as the observations you enter" (Aus-Res1) was widely shared. Participants suggested broadscale improvements in data were needed while also citing regional and jurisdictional differences in the type and quality of input data available. Areas such as Mallee and other, typically remote, landscapes were suggested as important focus areas for future research efforts.
2. Workforce and training: a lack of well-trained staff was also frequently cited as a key utilization challenge.
3. Inappropriate use of simulators or outputs: this relates to simulators being used by inexperienced users or for purposes that they were not designed for. While the immediate implications of poor utilisation include inaccurate outputs or poor decision making, participants suggested that a more significant implication is the likelihood of reputational damage.

Simulator development

Trade-offs and priorities

Participants viewed trade-offs as inevitable part of future development of fire simulators. The challenge for future development will be balancing the need to the fundamental aspects of simulators required for a strong base for simulator development while remaining open-minded to and supportive of new innovations.

Some frequently cited priorities for future fire simulator development included improving simulators capacity to better integrate dynamic connective feedback, spotting processes and meteorological information and effects (e.g. Coupled atmospheric fire modelling research). Some suggested focusing research efforts on improving coupled atmospheric fire modelling could have the greatest impact in terms of saving lies and minimizing losses during high-range fire events. Further, some stated that developing and testing simulators for use at finer spatial and temporal scales would improve their usefulness for prescribed burn planning.

Development challenges and gaps

Lack of inclusive and collaborative approaches to governance: many participants raised jurisdictional biases as a concern. Such biases can mean that newly updated models are only marginally better than existing simulators or are only endorsed for use in one jurisdiction. This was a minority perspective, strongly argued by some participants, particularly from smaller jurisdictions.

User engagement: Insufficient engagement with users at all stages of simulator design and development was cited by several participants as limiting the potential for fire simulators to meet use case needs and expectations. Views, however, were mixed on whether this was always a detriment to simulator development. Some participants suggested that the configurability of a modelling framework combined with availability of support can counteract issues around lack of engagement during development. Instead, they suggested that this was as an example of good fire development as it was efficient and streamlined, being unbound to the priorities and interests of all users.

Resourcing and governance issues

1. Short-term agency priorities do not align with long-term development needs: Several interviewees suggested this challenge could be partially overcome with a national, coordinated approach to identify current and future research needs and establishing long-term funding for simulator development.
2. Bureaucratic approval and workflow issues: Such issues prevent simulator developments being institutionalized or integrated with current systems. This perspective was most strongly shared by two participants from larger jurisdictions, one of whom described simulator development as "like watching paint dry" (NSW-SP2). This related to workflow issues wherein agencies place too much emphasis on centralised authority for decision making, which restricts access to simulators and opportunities for learning and skill development
3. Structural and cultural issues within fire agencies: some participants indicated that the prioritisation of systematic risk-based approaches to governance can be an obstacle to simulator development, creating a reluctance for agency management to support the development or utilisation of new simulators, versions or inputs.

Collaborative approach

Discussions of ideal future pathways and examples of “good” and “bad” simulator development in the past led many participants to argue for collaborative and coordinated approaches to research and development. There was a strong theme across interviews of the need for an inclusive approach to 1) build a communal and individual sense of ownership; and 2) to ensure development meets the needs of all jurisdictions. Views on the extent to which development should be nationally coordinated were varied and many participants shared examples of collaboration undermining effective or useful development.

Some participants put forward the need for a clear, nationally defined research and development agenda to focus investment on identified needs. Such an agenda could outline the development of a shared but flexible platform that can be tailored to the specific needs of different jurisdictions and landscapes. While most participants agreed that a one-size-fits-all approach is undesirable, a 'toolkit' approach to simulator development was suggested by many participants as the appropriate compromise. Some elements of development can or should be nationally coordinated, albeit with concerns remaining over governance and the process becoming overly bureaucratic.

Some of the stated benefits of such coordinated approach include streamlined investments and a larger pool of resources, diversity of inputs, skills and knowledge to bolster innovative capacity. Participants from smaller jurisdictions were particularly supportive of a national approach, noting the potential to leverage off larger budgets and resource capacity of larger states and opportunities for routinised training (significant during periods of no bushfire events), knowledge and skill transfer through nationally coordinated training. However, these participants also highlighted potential risks of failed development given relatively higher costs of buy-in and potential trade-offs for ownership and involvement in setting priorities for simulator development that align with state and territory needs.

Governance

Governance issues were widely cited and critiqued across interviews during discussions relating to most appropriate future pathways for research and development of fire simulators. As one participant stated, capturing the sentiments of others: “It's not just having a, a simulator that does stuff. It's the, the ecosystem if you like, around it and the context for how we're building simulators, which is really important” (Aus-Res2). As outlined above (see ‘Collaboration and coordinated/national approaches’), a strong theme across participants was the need to establish a clear long-term vision for future research and design of fire simulators. It was suggested that this required the following:

1. Establishment of a long-term leadership team with an established organizational ‘home’ to drive development, integration and utilization.
2. Clear commitments of long-term funding of research and development.
3. Pathways and/or guidelines to develop user skills and confidence using simulators (e.g. accredited training, evidence-based frameworks for decision-making, utilization guidelines, workflow integration and mentoring).
4. Flexibility in testing and use of simulators. “If you do let people have the freedom to try new things and play with it and make their own decisions about what's appropriate, then that'll give it the best bang for buck going forward” (SA-SP1).

Transparency and documentation

Transparency and documentation were widely referenced as features of good simulator development. Several respondents highlighted the importance of clearly documenting how and for what purposes simulators and subsequent versions have been developed, as well as assumptions and limitations. Interviewees provided examples of where clear documentation in development processes (or lack thereof) supports or hinders user understanding of modelling processes and capacity to interrogate the accuracy of outputs. These in turn, were described as critical for good utilisation through building user confidence and trust in simulators (as discussed above in ‘Reputation, trust and confidence’) Concerning outputs from an untransparent simulator, one user stated, “It could be really accurate, but I don’t know how it’s gotten there, so I can’t edit it to make it do what I want it to do… so even if it gave me an accurate result, I’m not going to use it because I don’t know it well enough for it to give me what I need.” (WA-TO1). Ultimately, users suggested that clear and thorough documentation supports good decision-making and accountability and bolsters decision-makers’ confidence.

Evaluation and verification

Several participants noted that evaluation and verification processes are the crux of what makes a good fire simulator. Perspectives varied, however, regarding what form these processes must take to be appropriate, tangible and useful in various contexts. The value of scientific/peer-reviewed validation proved to be a significant point of difference among interviewees. For example, research use case interviewees emphasized the critical importance of scientific, peer-review processes for developing evidence-based frameworks for simulator utilization and decision-making. It was suggested by some participants that having clear and 'evidence-based frameworks' is critical for ensuring accountability and 'safer' conditions for decision-making. As one interviewee stated "It's much, much easier and safer in a way from a legal sense to have an output…from a system that's been tested and gone through a peer-review process and scientifically established rather than making critical decisions from a briefing from an individual, which is somewhat more subjective" (Aus-Res2).

Others felt that while scientific validation serves a foundational role in simulator development, the priority should be developing a nationally accepted approach to verification and testing. Some interviewees had observed a broad shift in fire management and research towards acknowledging the limits of scientific evidence for good decision making in operational contexts, suggesting that historical simulator development has sometimes "…diverged away from that purely scientific basis to try and get something that works" (Aus-SP1). One participant indicated, for example, that having a clear statement of purpose for models, "even different aspects of a model depending on what it's designed to do" [US-Res1] would be helpful for developing clear and differentiated evaluation frameworks.

Many participants highlighted the importance of developing simulators with transparent and configurable modelling frameworks so that evaluation can happen through testing in real-time (allowing users to first understand processes behind outputs and subsequently, alter inputs or interrogate outputs depending on real-time observations). According to QLD-SP1, a good fire simulator allows users to "play around with the inputs and having that user interface or user operations where, where people can test sensitivity in real time, test it to sensitivity of the outputs and understand why you are getting a particular result is really important".

Workshop 1

Findings from Workshop 1 indicated that establishing discrete standards for individual criteria for simulator development and use would be a significant challenge. This was primarily because when participants were asked about their expectations of simulators and their willingness to adopt changes to simulators, participants consistently stated that it depended on the context. What they expected from simulators and what they valued most about good simulators varied according to both the broader circumstances of the user (e.g. the organisation, jurisdiction, use case, experience level) and the context of any given use (e.g. type, size and risk of fire, purpose of the simulation, audience).

Ultimately, a whole variety of factors were noted as being significantly important to good simulators. Of the list of 35 criteria identified as important, when participants were asked to provide a rating out of 5, the average rating for each of the criteria was 4.0 (standard deviation 0.30). In addition, participants also frequently noted the interconnectedness between the different factors, furthering the challenge of creating discrete benchmarks for individual criteria.

As was observed in the interviews, when asked about what aspects of simulator development should be prioritised, participants called for attention, planning and resourcing not just on technical improvements but also on broader social, contextual factors that affect successful simulator development and use. Some considered these contextual factors equally if not more important than technical ones and necessarily integrated into the simulator design process. Per one participant;

"...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, knowing that we're not going to come up with a perfect simulator anytime soon. And I think it's important to build that into the thinking process rather than have an add-on at the end." Workshop 1 participant

There was some variation in the most important performance criteria between the Tactical, Strategic and Research use cases (Table 6). However, all groups rated 'Good reputation: trust in underlying research' in their top five. Tactical users frequently prioritised tools that were simple, easy to use and that can quickly produce an output that the user and the audience would trust. Both Strategic and Research users had more of a focus on building complexity and nuance into the simulation process. Both prioritised configurability of inputs and the ability to compare a variety aspects of fire behaviour or outcomes simultaneously.

Across all use cases, on average, the bottom five (least important) criteria were *good price*, *configurable* (Useable by different user types and skill levels), *fast* (speed of specific part of process e.g. fire spread simulation), *good software framework* (open source) and *fast* (speed of entire process, from data input to usable output).

TABLE 6 - MOST IMPORTANT PERFORMANCE CRITERIA BY USE CASE, FROM WORKSHOP 1

GROUP	MOST IMPORTANT CRITERIA
Tactical	Well supported: access to training and education
	Fast: speed of specific part of process
	Easy to use: including interface
	Good input process: ease of preparation
	Good reputation: trust in underlying research
Strategic	Good outputs: able to be audited
	Good outputs: range of detailed spatial outputs
	Good reputation: trust in underlying research
	Configurable: Tweakability, tuneability, tailorability
	Good modelling framework: able to run ensembles
Research	Good modelling framework: consider uncertainty
	Configurable: Tweakability, tuneability, tailorability
	Specialised/universal: Versatile - able to model all aspects of fire behaviour
	Specialised/universal: Targeted – specialises in specific aspects of fire behaviour
	Good reputation: trust in underlying research

When considering a change in simulator usage, participants agreed that trade-offs between important simulator criteria were inevitable but were generally reluctant to accept any decrease in performance accompanying a change in simulator usage. The most important criteria to support a change were validation, reputation, support and compatibility. This finding was consistent across use cases and across the nature of change (input layer, new model version, new model).

Workshop 2

Workshop 2 was designed, in part, for participants to identify the thresholds at which simulators become fit or unfit for purpose. It also encouraged participants to paint a picture of what simulators look like beyond those thresholds, (i.e. what truly bad and good simulators look like for them). The first two scenarios asked participants to respond within the context of their own simulator use. Scenario 3 encouraged participants to think beyond the technical, social and financial constraints of the simulators they currently use and articulate what simulators would ideally be capable of if they were to have the greatest improvement on outcomes (e.g. new capabilities or the possible role of new technologies such machine learning or artificial intelligence). Much like in Workshop 1, there were productive discussions and ideas raised throughout the session, however, identifying clear or consistent benchmarks and priorities proved difficult. These difficulties further indicated the challenges of generalising about the needs of simulator users and creating widely applicable definitions of good simulators.

In Scenario 1, designed to elicit clear benchmarks of unacceptable simulator performance, participants often found it easier to discuss minimum *acceptable* performance than to identify what *unacceptable* performance looked like. Participants were typically able to speak in general terms when articulating what they wanted from a good simulator. They were often unable, however, to attach tangible, numeric or objective benchmarks to their needs. They struggled to benchmark their own simulator requirements and even more so, were not easily able to generalise about benchmarks that represented the broader simulator community, often saying ‘it depends’. Technical performance criteria (e.g. accuracy, speed, resolution and ensembles) were the exception to this. A small number of participants were able to suggest measurements for these criteria and propose numeric values, however there was no wide consensus among users.

These challenges of articulating measurable values were consistent across all three scenarios. Scenarios 2 and 3 both asked users about the improvements needed to ensure their willingness to change. Scenario 2 asked about the minimum participants would expect to support and change and Scenario 3 asked about ideal improvements, or game-changers. For participants, making changes to a simulator were, again, generally stated as the presence or absence, increase or decrease, of a specific feature that related to one of the criteria. They put forward a wide range of attributes and features that were important for their uses (see Discussion for a summary of participant suggestions for simulator capabilities, availability and quality of data and governance).

Responses were, again, rarely framed as measurable values independent of context, but were instead framed within the context of a participant’s own use. The importance that participants gave to criteria typically related to how well that criteria worked for them in their circumstances. For example, participants often stated that users would be unwilling to accept any decrease in performance, particularly in accuracy and this would therefore vary from user to user depending upon how well a simulator performed in their context. There appeared to be a jurisdictional and use case aspect to perceived simulator performance. Participants, frequently Tactical users, noted that different jurisdictions can have different simulator needs, such as input data or models (e.g. vegetation models) and may require varying levels of support to maintain simulators or switch to new options. Some participants commented, for example, that Phoenix has been tailored to Victoria and New South Wales and may be more accurate than Spark in the majority of use contexts (particularly dry eucalypt forests). Such a reality may result in different levels of motivation for users Victoria or New South Wales to support a switch compared with other states.

Questionnaire

Participant demographics

Most respondents were from Victoria (33%) and New South Wales (25%), followed by Queensland (12%) and Tasmania (9%). Only a few participants represented each of Western Australia, South Australia, the Northern Territory and the Australian Capital Territory. Most participants worked in metropolitan areas (46% in capital cities and 16% in other metropolitan cities), about one third in regional areas (28%) and the smallest proportion in remote areas (9%). The vast majority of participants worked for government departments or agencies (78%). The remaining participants came from research organisations (12%), private consultancies (4%) or other organisations (6%).

Note that, with a total sample size of 67, very low percentage responses (e.g. less than 5%) are less meaningful as they represent 4 or fewer people.

The average age of participants was 48 years old. 79% of all participants had at least 5 years of experience; almost half (49%) had more than 10 years of experience. On average, participants spent about one quarter of their role using fire simulators, however answers varied widely, ranging from 1-90%. Phoenix (84%) and Spark (61%) were the most common simulators, but most participants (63%) used multiple simulators in their role.

The primary use case of most participants was Tactical operations (55%), followed by Strategic planning (25%) and Research (15%). Five per cent stated other uses. The majority of participants (92% overall) stated that they used simulators for activities that spanned more than one use case. However, this figure was only 44% amongst researchers. Across all users, the most common uses for fire simulators were operational predictions (real-time) (62%), scenario training (55%), fire spread reconstruction (53%), landscape risk assessment (45%), risk mitigation and treatment planning (43%) and operational suppression advice (41%).

Most participants provided simulator output for audiences with basic (49%) or intermediate (30%) knowledge of simulators. It was rare for audiences to have either zero or expert knowledge. These results were consistent across use cases.

User requirements

Fire characterisation accuracy benchmarks

Participant requirements for accuracy of fire characterisation were somewhat consistent across the four factors, rate of spread, area burned, fire behaviour and impacts (Table 7). Fire behaviour relates to the characterisation of factors such as intensity, flame height and ember transfer. Impacts relate to the characterisation of consequences for values such as communities, the environment and asset loss. The median participant response was of at least 60% accuracy and a maximum of 85-90% across all the four variables, although acceptable minimum accuracy was as low as 10% for some users. Achieving accuracy beyond 90% was generally considered superfluous to achieving greater outcomes. The boxplots in Figure 1 below show variability between the four different use groups. Whilst there was some variability across the use cases, participant responses were generally consistent, regardless of their primary use case (see Appendix 1 for medians and ranges specific to each of the different use cases).

TABLE 7 - BENCHMARKS FOR FIRE CHARACTERISATION ACCURACY

Variable	Minimum accuracy required (%)		Maximum accuracy required (%)	
	Median	Range	Median	Range
Rate of spread	60	10-95	90	50-100
Area burned	60	10-90	90	40-100
Fire behaviour	60	20-90	85	30-100
Impacts	60	10-90	87	49-100

Most users identified accuracy of rate of spread as the most important out of the four fire characterisation factors,

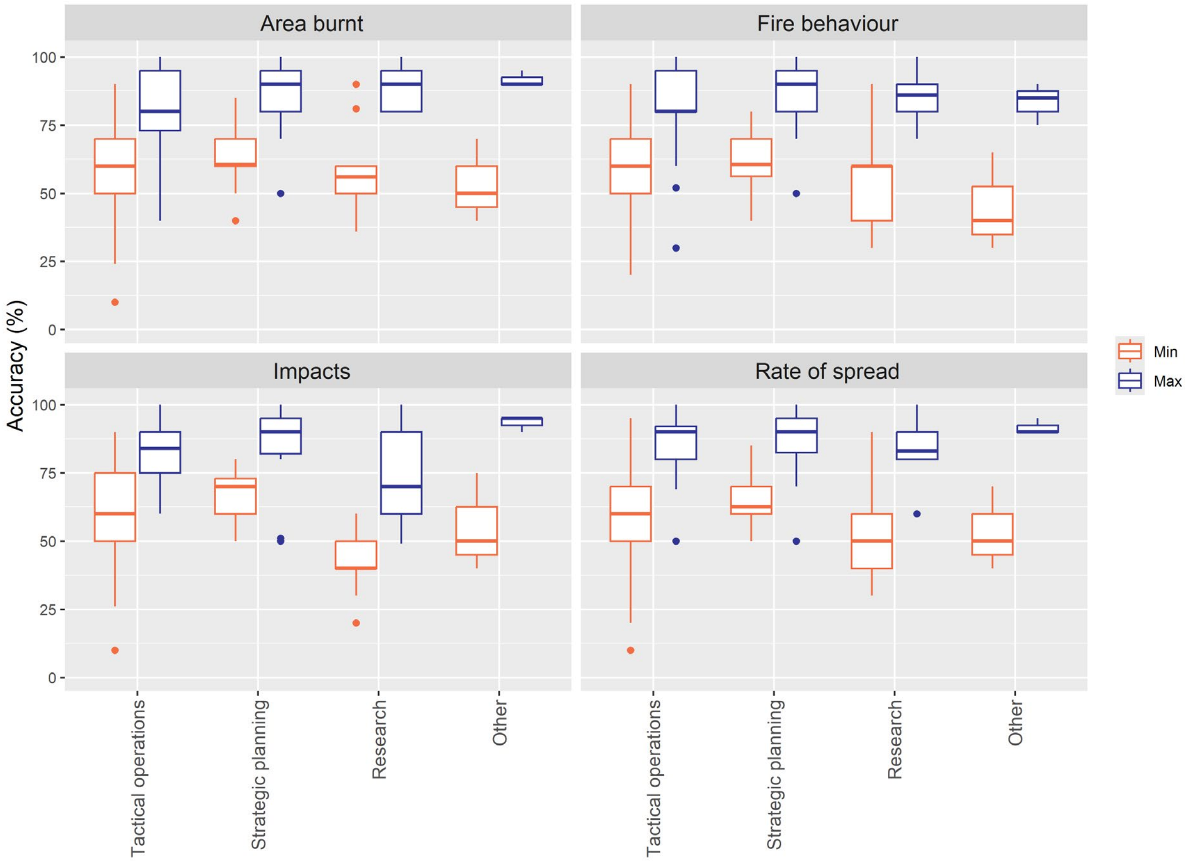

FIGURE 1 - PERFORMANCE BENCHMARKS FOR FOUR ASPECTS OF FIRE CHARACTERISATION: AREA BURNT, FIRE BEHAVIOUR, IMPACTS AND RATE OF SPREAD. ORANGE BOXPLOTS SHOW MINIMUM REQUIREMENTS, BLUE BOXPLOTS SHOW MAXIMUM. RESULTS ARE SEPARATED ACCORDING TO THE FOUR SIMULATOR USE CASES. THE LOWER AND UPPER HINGES CORRESPOND TO THE FIRST AND THIRD QUARTILES. WHISKERS EXTEND UP TO 1.5 X THE INTERQUARTILE RANGE.

participants were broken down according to their primary use case. 57% of all users ranked fire behaviour either first or second most important (Figure 2).

Looking at how often each criterion was ranked in the top two however, both Tactical and Strategic ranked rate of spread as the most important with Fire behaviour as the second most (Figure 3 and Figure 4). For Research, however, Area burned was just as important as Rate of spread, with Fire behaviour the third most important (Figure 5).
Rate of spread was widely considered the most important aspect of fire to characterise, regardless of use case. There was variability and some disagreement between use cases for other three aspects. Area burnt was the least frequently ranked top two for Tactical. In contrast, researchers ranked this factor in their top two as frequently as they did rate of spread. Fire behaviour was consistently important among users and was ranked in the top two by over 50% of participants for each use case. Accurate characterisation of impacts was relatively unimportant to Strategic and Research users, but more important to Tactical users.

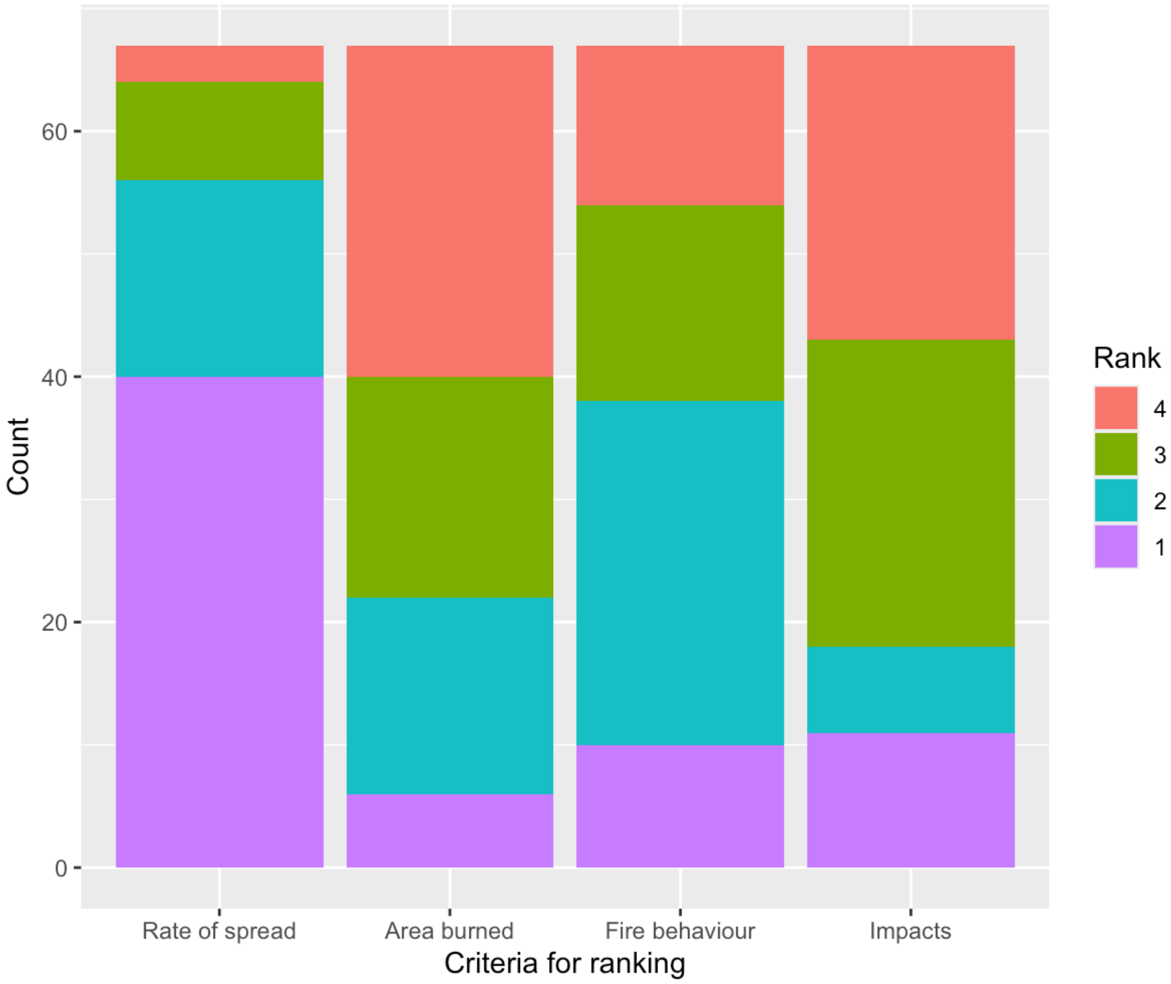

FIGURE 2 - USER RANKINGS OF THE IMPORTANCE OF FOUR MEASURES OF FIRE SIMULATOR PERFORMANCE; RATE OF SPREAD, AREA BURNED, FIRE BEHAVIOUR AND IMPACTS.

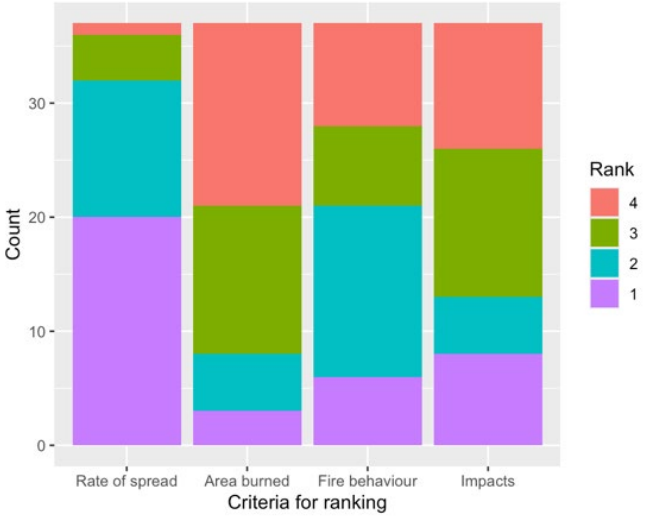

FIGURE 3 -TACTICAL USER RANKINGS OF THE IMPORTANCE OF FOUR MEASURES OF FIRE SIMULATOR PERFORMANCE; RATE OF SPREAD, AREA BURNED, FIRE BEHAVIOUR AND IMPACTS.

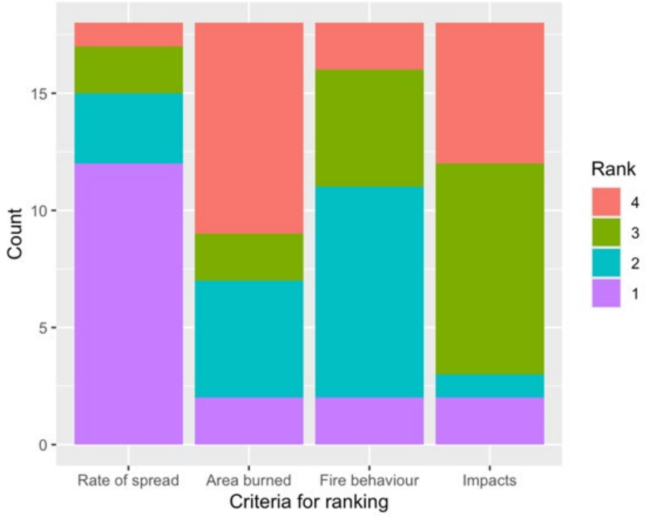

FIGURE 4 - STRATEGIC USER RANKINGS OF THE IMPORTANCE OF FOUR MEASURES OF FIRE SIMULATOR PERFORMANCE; RATE OF SPREAD, AREA BURNED, FIRE BEHAVIOUR AND IMPACTS.

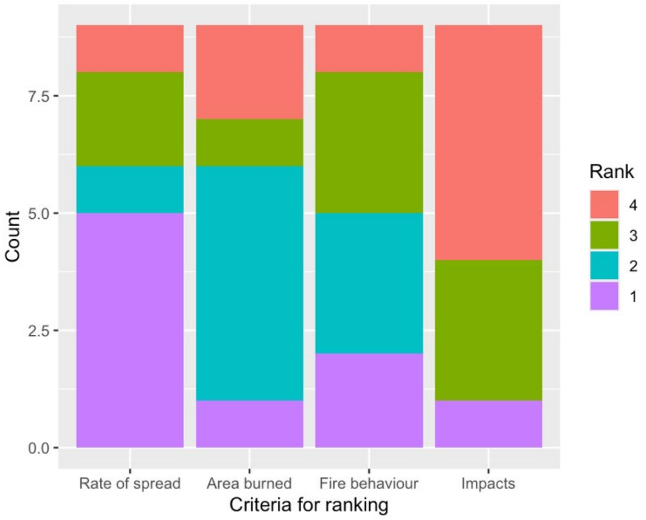

FIGURE 5 - RESEARCH USER RANKINGS OF THE IMPORTANCE OF FOUR MEASURES OF FIRE SIMULATOR PERFORMANCE; RATE OF SPREAD, AREA BURNED, FIRE BEHAVIOUR AND IMPACTS.

Performance benchmarks

Benchmarks for four criteria related to technical simulator performance are shown in Table 8. These are spatial resolution, model timestep, output speed and ensemble capacity. The minimum requirements of users (also shown in table 8) varied widely. Requirements did not obviously vary by use case (Figure 6). Appendix 2 shows the boxplots with untransformed axes, with and without outliers (Figure 13 and Figure 14) to understand the scale of variability between participant responses.

TABLE 8 - BENCHMARKS FOR SIMULATOR PERFORMANCE FEATURES (MINIMUM AND MAXIMUM)

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
Model timestep[4]	10 minutes	60 minutes	1 second – 3 hours	1 minute- 24 hours
Output speed	1 minutes	15 minutes	1 second – 5 hours	30 seconds - 2 weeks
Number of ensembles	1	50	0 – 20,000 ensembles	3 – 20 million ensembles

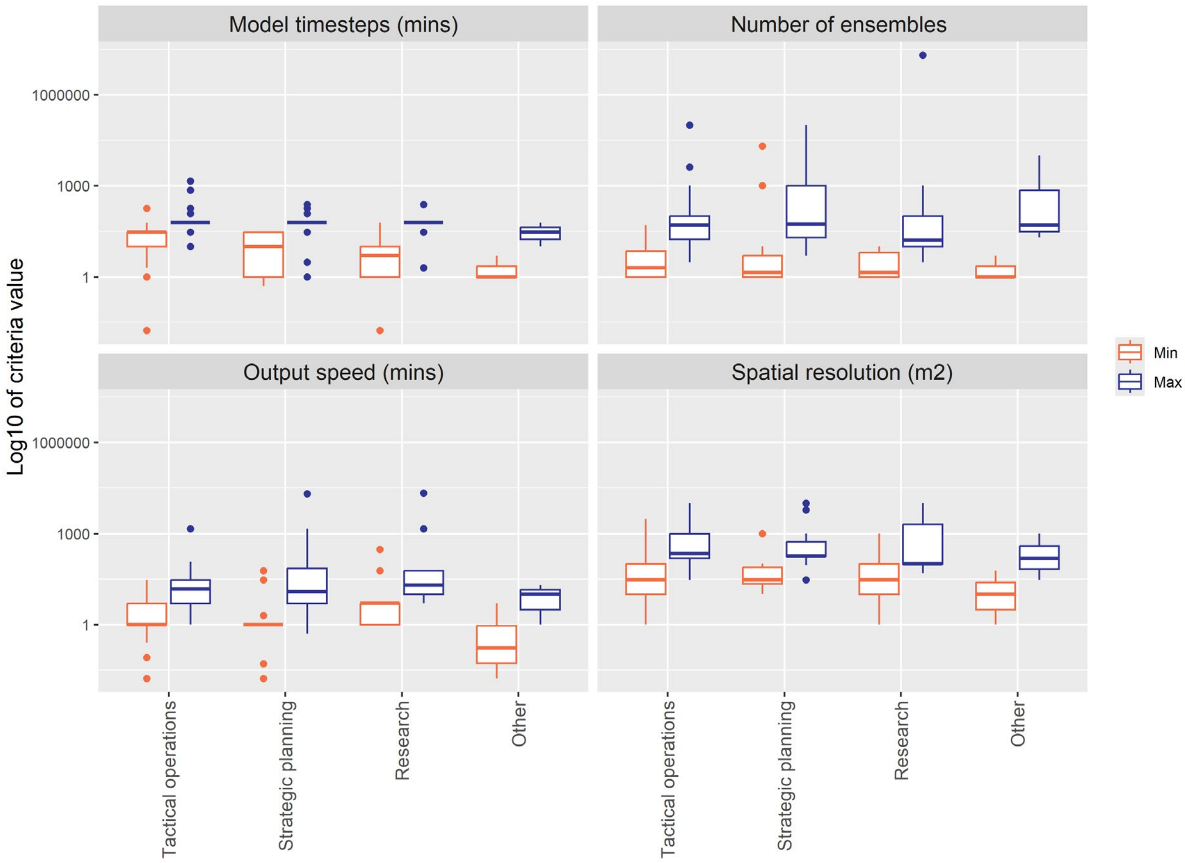

FIGURE 6 – PERFORMANCE BENCHMARKS FOR FOUR ASPECTS OF SIMULATOR TOOLS; MODEL TIMESTEPS, NUMBER OF ENSEMBLES, OUTPUT SPEED AND SPATIAL RESOLUTION. LOG10 TRANSFORMATION OF CRITERIA VALUE. ORANGE BOXPLOTS SHOW MINIMUM REQUIREMENTS, BLUE BOXPLOTS SHOW MAXIMUM. RESULTS ARE SEPARATED BY USE CASE. THE LOWER AND UPPER HINGES CORRESPOND TO THE FIRST AND THIRD QUARTILES. WHISKERS EXTEND UP TO 1.5 X THE INTERQUARTILE RANGE.

[4] This question was intended to capture the timestep of potential outputs. It is possible that some respondents were referring to an internal timestep rather than the timestep of outputs. Results should be interpreted with this caveat in mind.

User preferences

Section C of the questionnaire explored user preferences. It asked participants about the most important factors that contributed to five different areas deemed significant for good simulators; user experience (high and low quality), simulator outputs (good and bad), simulator use outcomes (good and bad), simulator trustworthiness (according to the user and to the audience) and the simulator ecosystem (for the user organisation and for Australia).

Participants were asked a set of two corresponding questions for each of the five topics to understand which simulator factors were seen to have the greatest effect on simulator use and whether these factors had both positive and/or negative effects. Each of the figures (7-11) contains two graphs that focus on both the positive (left) and negative (right) effects for each of the topics. Each graph shows a variety of different variables related to one of the five areas of simulator use or development (x-axis) and the proportion of participants from each of the four use cases who ranked each variable in their top five most important (y-axis). In most cases the same features that drove a positive effect also drove a negative effect. There were, however, some examples where results differed. These findings are discussed in detail below.

Differences between use cases

Whilst there is some variability between use cases the priorities of different users are generally consistent.
Note: Research (9) and Other (3) had far fewer participants than Technical (37) and Strategic (18). Certain results, such as overall user preferences, are therefore skewed towards the responses of technical and, to a lesser extent, Strategic users. 'Other' should not be seen as representative of a particular set group. It was none-the-less included in these graphs to ensure the accurate reflection of the diversity across participants.

User experience

Among all participants, Stable platform (94%), Adjustable settings (88%), Intuitive Graphical User Interface (75%), Easy to spot and fix errors (60%) and Quick to set up (54%) were most commonly in the top five most important factors contributing to a high-quality user experience. Similar values were observed for a low-quality user experience; however, poor tech support was also selected in the top five by over 50% of participants (58%).

Stable platform was not only the most common factor to appear in the top five but it was also most frequently ranked as the number one driver of both high-quality (35% of users) and low-quality user experience (45% of users).

Figure 7 below shows considerable consistency in responses between use cases. However, a few instances of variability occurred.

Key differences between use cases

High-quality user experience

- Stable platform was slightly less frequently important for Research than Tactical or Strategic.
- Intuitive Graphical User Interface was more frequently important to Research than Tactical or Strategic

Low-quality user experience

Adjustable settings were more frequently important for Research than Tactical and Strategic

- Quick setup was more frequently ranked important for Tactical than Strategic or Research.
- Standalone app was more frequently ranked important for Tactical and Research than Strategic.
- Tech support was more frequently ranked important for Strategic than Tactical and Research

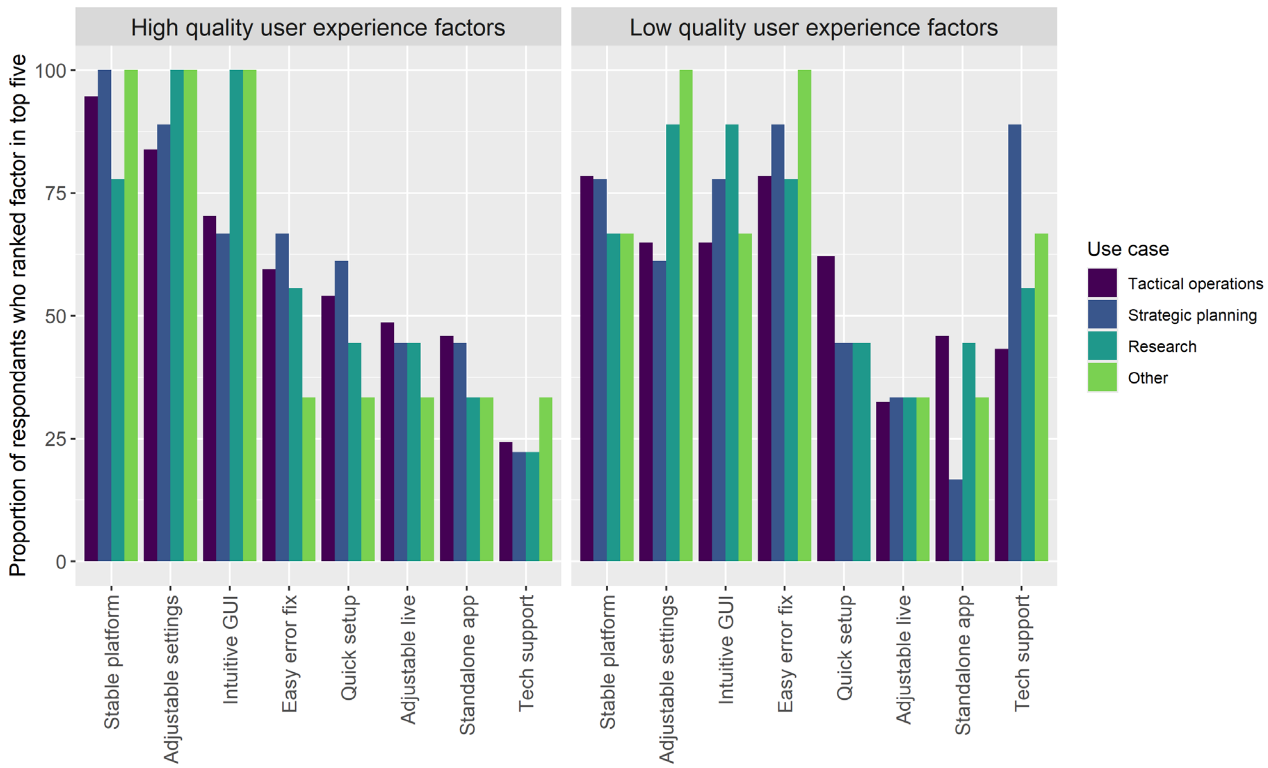

FIGURE 7 – USER PRIORITISATION OF FACTORS THAT AFFECT USER EXPERIENCE. BAR GRAPHS SHOW THE PROPORTION OF PARTICIPANTS (Y-AXIS) WHO SELECTED EACH FACTOR (X-AXIS) IN THEIR TOP FIVE MOST IMPORTANT FACTORS FOR INFLUENCING HIGH-QUALITY USER EXPERIENCE (LEFT GRAPH) AND LOW-QUALITY USER EXPERIENCE (RIGHT GRAPH). USERS WERE SEPARATED ACCORDING TO THEIR PRIMARY USE CASE.

Output quality

The top three most important factors contributing to both high-quality simulator outputs and low-quality simulator outputs were the same three factors, which all related to input data. These were fuel data (88%), weather data (87%) and mapped location (67%). Mapped location, however, was disproportionately important for Tactical users. Factors relating to the simulator itself and the knowledge and skills of stakeholders were all ranked lower than the data factors. The most ranked number one for 'High quality outputs' was weather data (21%) and for 'low quality outputs' was fuel data (21%). There was much agreement between the use cases about the importance of factors contributing to output quality, however, there were some differences too (Figure 8).

Key differences between use cases

High-quality outputs

- Good mapped location was more frequently important for Tactical than Strategic or Research.
- Good spatial resolution was slightly less important for Tactical than Strategic and Research
- Fast outputs were less frequently considered important for Research than Tactical and Strategic users.

Low-quality outputs

- Good weather data was more frequently important for Research than Tactical and Strategic.
- Good mapped location was less frequently ranked as important for Strategic users than Tactical and Research.
- Fast outputs were less frequently ranked important for Research than Tactical and Strategic.
- Skilled operator was more commonly seen as important for Strategic than Tactical or Research.

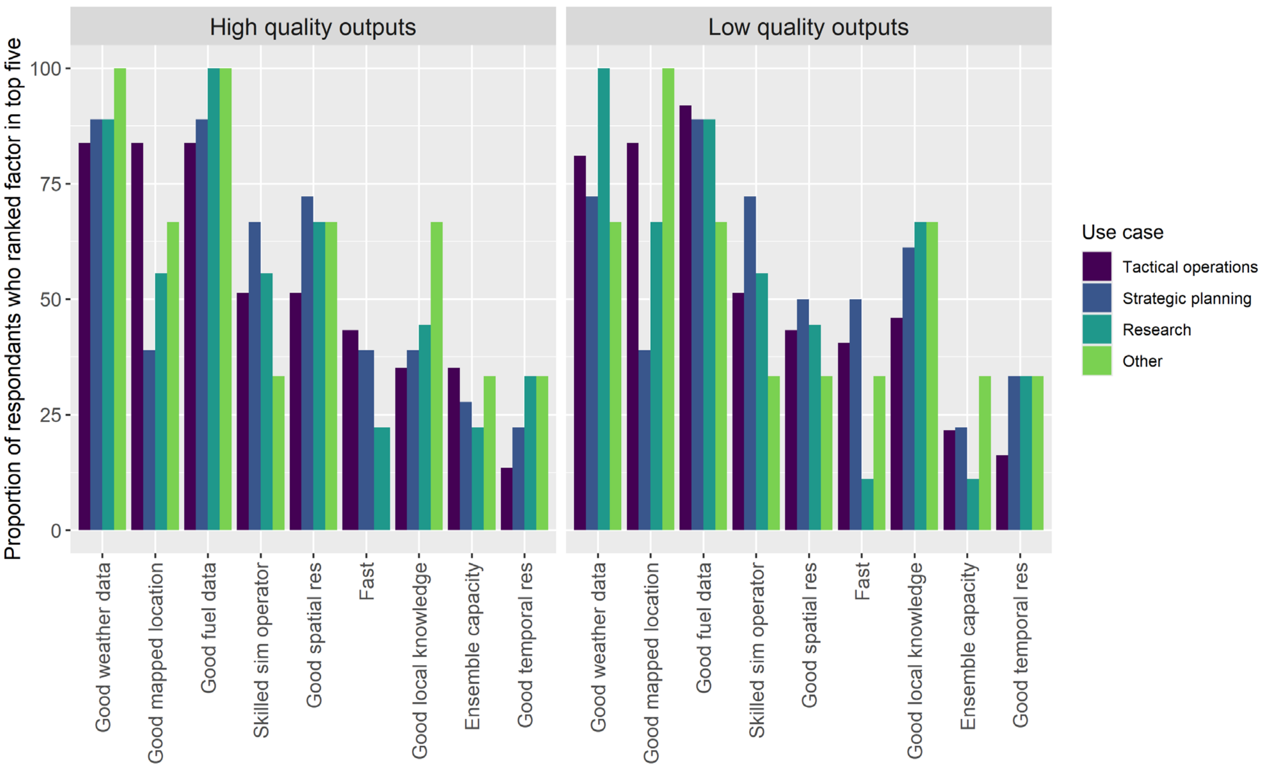

FIGURE 8 - USER PRIORITISATION OF FACTORS THAT AFFECT SIMULATOR OUTPUTS. BAR GRAPHS SHOW THE PROPORTION OF PARTICIPANTS (Y-AXIS) WHO SELECTED EACH FACTOR (X-AXIS) IN THEIR TOP FIVE MOST IMPORTANT FACTORS FOR INFLUENCING HIGH-QUALITY OUTPUTS (LEFT GRAPH) AND LOW-QUALITY OUTPUTS (RIGHT GRAPH). USERS WERE SEPARATED ACCORDING TO THEIR PRIMARY USE CASE.

Outcome quality

Communication of assumptions came up as the most important factor determining both good and bad outcomes of simulator use. The factors most frequently ranked in participants top five for high-quality simulator outcomes were Communication of assumptions (69%), Quality of the visualisation (61.2%), Simplicity of the overall output presentation document (58.2%), Face-to-face briefing/presentation of outputs (57%) and Audience trust in simulator user (52%). Simplicity of the overall output presentation document was most frequently ranked first by users (19%).

Similarly for low-quality simulator outcomes, the most important factors were Communication of assumptions (75%), Face-to-face briefing/presentation of outputs (63%), Complexity of the overall output presentation document (58%), Quality of the visualisation (52%) and Audience trust in simulator user (52%). Quality of the visualisation was most often ranked number one by users (21%). Once again, participant preferences were generally consistent across the use cases as can be seen in Figure 9, although some differences are outlined below.

Key differences between use cases

High-quality outcomes

- Audience trust in both operator and tool was more important for Tactical than Strategic and Research.
- Detail of visualisation more important for Research and Strategic than Tactical.
- Comparison of ensembles was important for Strategic and Research but not for Tactical.

Low-quality outcomes

- Audience trust in operator was important for Tactical but far less for Strategic and Research
- Experienced audience was less important for Strategic than Tactical and Research.
- Detail of visualisation was important for Strategic but less so for Tactical and Research.

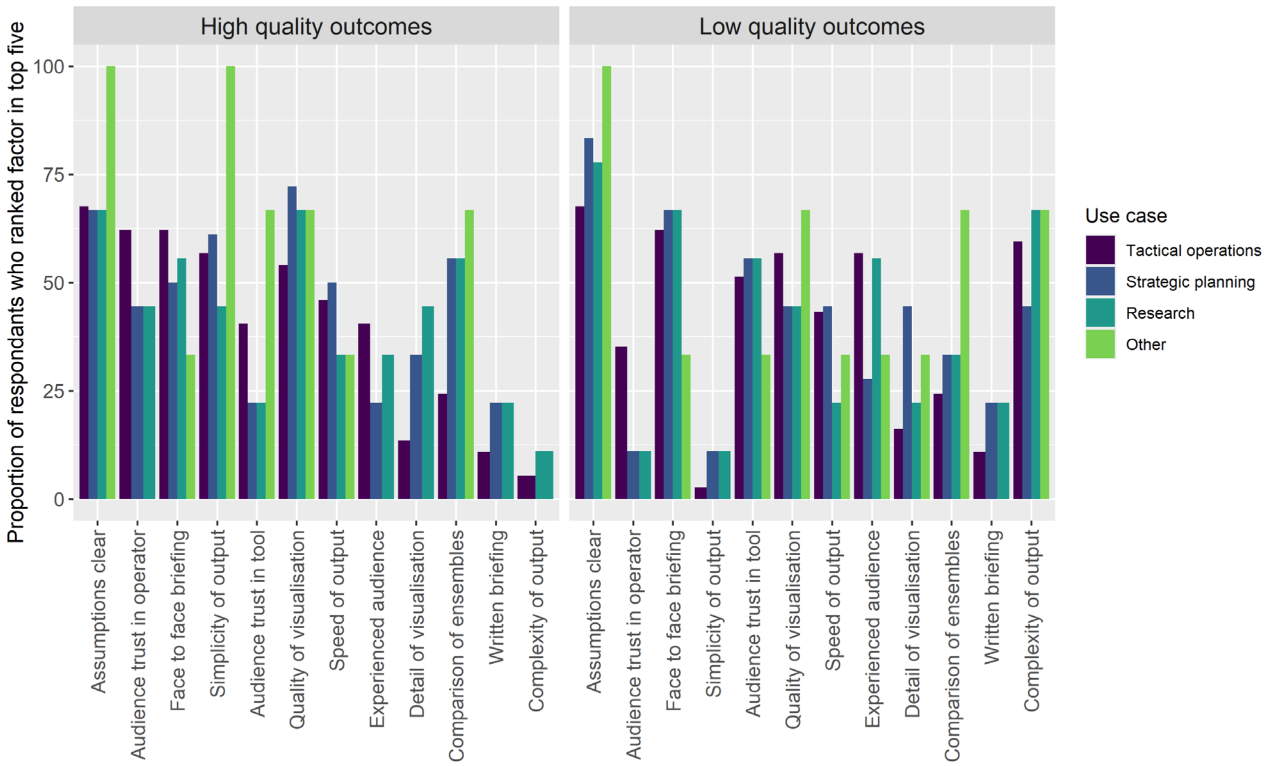

FIGURE 9 - USER PRIORITISATION OF FACTORS THAT AFFECT OUTCOMES FROM FIRE SIMULATORS. BAR GRAPHS SHOW THE PROPORTION OF PARTICIPANTS (Y-AXIS) WHO SELECTED EACH FACTOR (X-AXIS) IN THEIR TOP FIVE MOST IMPORTANT FACTORS FOR INFLUENCING HIGH-QUALITY OUTCOMES (LEFT GRAPH) AND LOW-QUALITY OUTCOMES (RIGHT GRAPH). USERS WERE SEPARATED ACCORDING TO THEIR PRIMARY USE CASE.

Trustworthiness

Two questions were used to understand participant perspectives on the trustworthiness of simulators. The first question asked about the most important factors that affected a user's own trust in a simulator, whilst the second question asked what factors they assumed most affected audiences' trust in a simulator. There was some notable variability in responses to the two questions. Accuracy of outputs was the most common response for both users (75%) and their audiences (81%). The remaining top four varied between the two questions. For user's trust these were validation (70%), clarity and assumptions included (57%), peer-reviewed science (55%) and platform stability (45%). For assumed audiences' trust, the remaining top four were organisational endorsement (57%), user reputation (55%), clarity including assumptions (58%) and platform reputation (52%). Whilst these findings provide insight into users' perceptions of their audiences', further research of audiences is necessary to validate whether these findings reflect audiences' actual needs.

The popularity of each of the factors were largely consistent across the use cases. Some variability can be seen for several factors outlined below (Figure 10).

Key differences between use cases

User trust in a simulator

- Platform stability was more frequently ranked important for Tactical users than Strategic and Research
- Peer-reviewed science and documentation were more frequently ranked important for Research and less so for Strategic then Tactical.
- Factors related to reputation (user reputation, sub-model reputation, developer and platform reputation) were typically less important for Research than Tactical and Strategic.

Audience trust in a simulator

- Validation was more likely ranked important for Research but less so for Tactical and Strategic.
- User (operator) reputation was more important for Tactical and Strategic than Research.
- Platform reputation was more important for Strategic than Tactical or Research.

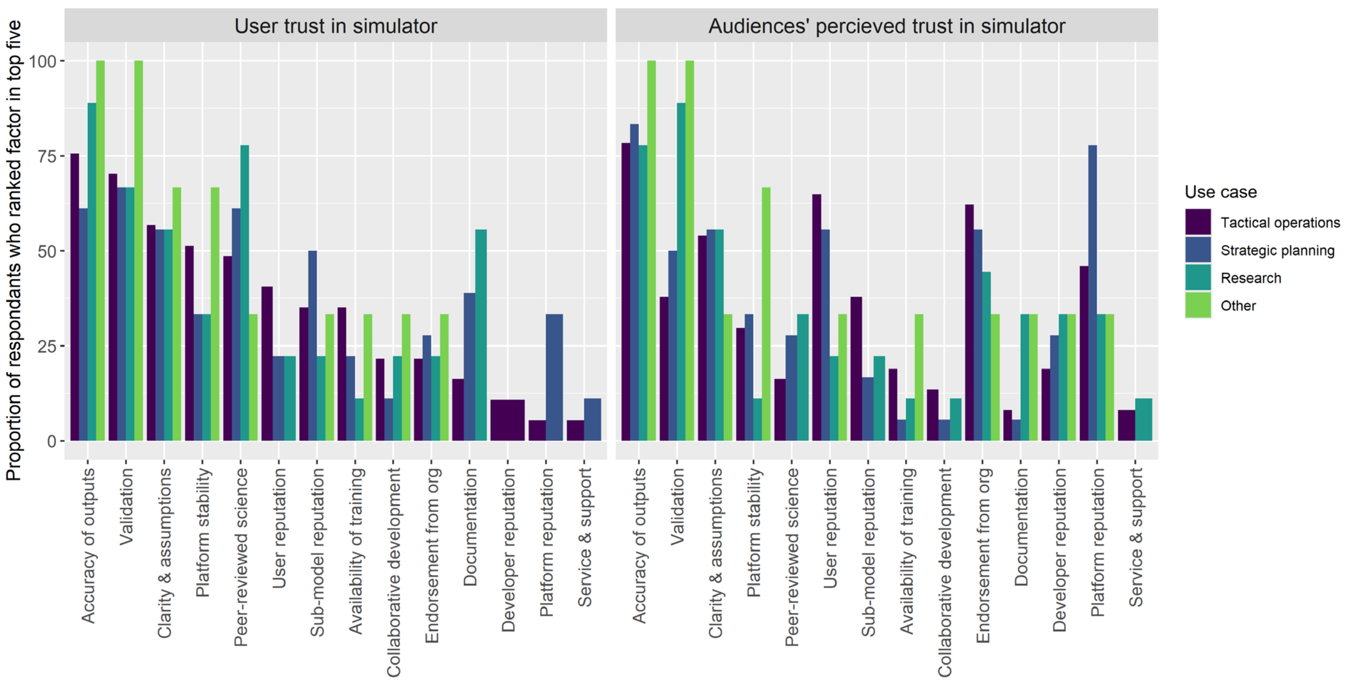

FIGURE 10 - USER PRIORITISATION OF FACTORS THAT AFFECT USER AND AUDIENCE TRUST IN FIRE SIMULATORS. BAR GRAPHS SHOW THE PROPORTION OF PARTICIPANTS (Y-AXIS) WHO SELECTED EACH FACTOR (X-AXIS) IN THEIR TOP FIVE MOST IMPORTANT FACTORS FOR INFLUENCING USER/OPERATOR TRUST IN A FIRE SIMULATOR (LEFT GRAPH) AND THEIR AUDIENCES' TRUST IN A SIMULATOR (RIGHT GRAPH). USERS WERE SEPARATED ACCORDING TO THEIR PRIMARY USE CASE.

Simulator ecosystem

Participants were asked which aspects of the broader simulator ecosystem were most important. Importance was considered with respect to national priorities and priorities for the participant's organisation. The top two selected by participants for organisational level were Improved fire behaviour models (66%) and Improved validation of data (64%). National priorities had the same top two but in a different order; Improved validation of data was the most popular (66%) and Improved fire behaviour models (60%) the second. The remaining three top five factors varied between the two questions. For organisation these were Compatibility with existing mapping (52%), Training for users (51%) and Compatibility with existing systems (51%). For National they were Improved transparency & documentation (58%), Improved governance (52%) and Improved inputs - data collection & storage (49%). Figure 11 below shows that again there was some variability between the use cases, but a great deal of consistency for most variables.

Key differences between use cases

Organisational level contributors to simulator ecosystem

- User training was less frequently ranked important for Strategic than Tactical or Research.
- Compatibility with systems was more frequently ranked important for Tactical than Strategic and never ranked top five for Research.

-

National level contributors to simulator ecosystem

- Fire behaviour models were less frequently ranked important for Tactical than Strategic and Research.
- Inputs were often ranked as important for Strategic and Research and less so for Tactical.
- User certification was more commonly ranked as important for Tactical than for Strategic and Research.

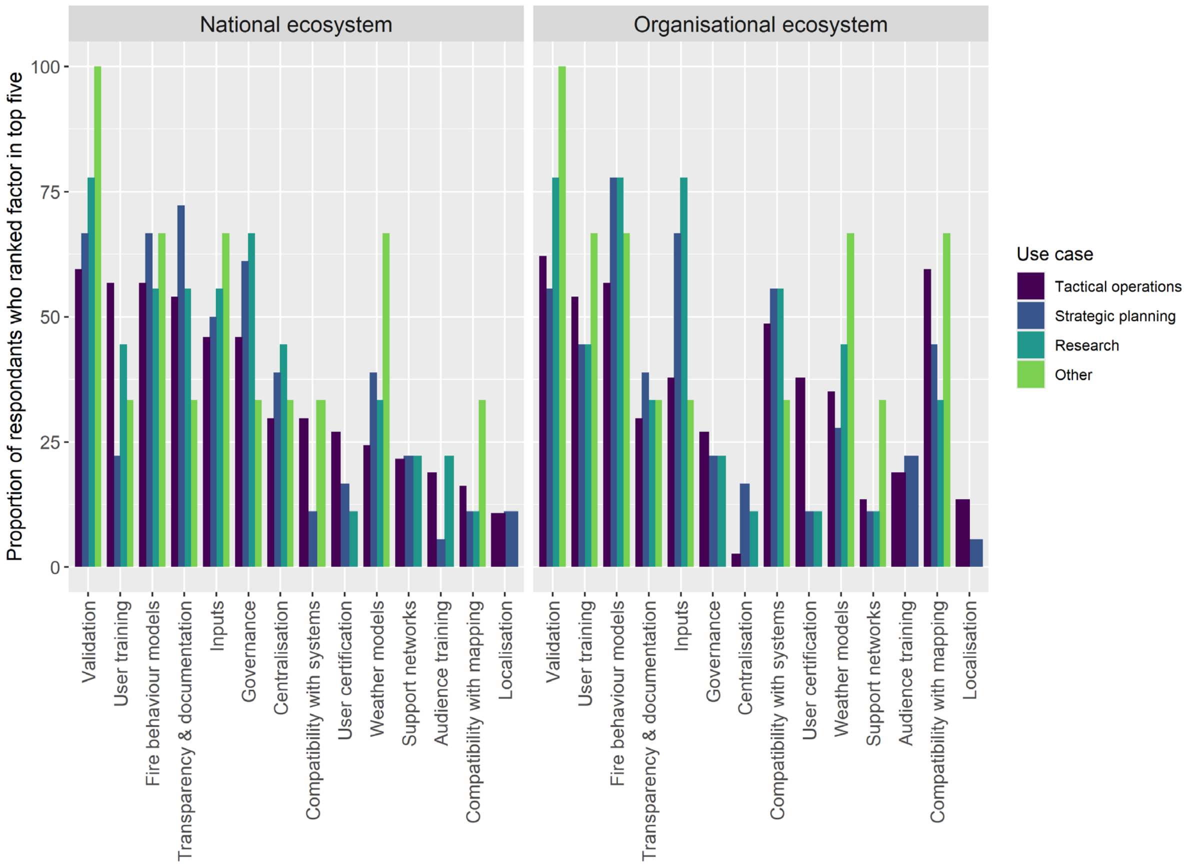

FIGURE 11 - USER PRIORITISATION OF FACTORS THROUGHOUT THE BROADER SIMULATOR ECOSYSTEM THAT ARE MOST RELEVANT TO THE DEVELOPMENT AND MAINTENANCE OF GOOD SIMULATORS. BAR GRAPHS SHOW THE PROPORTION OF PARTICIPANTS (Y-AXIS) WHO SELECTED EACH FACTOR (X-AXIS) IN THEIR TOP FIVE MOST IMPORTANT FACTORS. FACTORS MOST IMPORTANT AT A NATIONAL SCALE WERE CONSIDERED IN THE LEFT GRAPH. FACTORS RELEVANT AT THE PARTICIPANT'S ORGANISATIONAL LEVEL WERE CONSIDERED IN THE RIGHT GRAPH. USERS WERE SEPARATED ACCORDING TO THEIR PRIMARY USE CASE.

Discussion

The following discussion considers how findings from across the interviews, workshops and questionnaire can support the project's objective of guiding development of fire simulators to benefit all users in the future. We found that simulator user needs are diverse, context dependent, evolving and resistant to simplification. Two recurring themes were, firstly, the importance of ensuring simulator development is cognisant of social factors and not just technical ones and, secondly, that a robust research foundation, relationship building and proven reliability and quality are all important for creating legitimacy and support for simulators and their outputs. We also found a set of criteria, which were identified as consistently relevant to ensuring fit-for-purpose fire simulators.

There were some consistent differences between the main use cases (*Tactical*, *Strategic* and *Research*) but perhaps surprisingly, use case was often a weak predictor of the relative importance of key criteria. Jurisdiction and geography are important drivers of fire simulator user needs, given Australia's strong regional variation in fire regime, risk and organisational resourcing. Despite – or perhaps because of – the considerable body of insights generated from user engagement; it was challenging to crystallise user requirements in the form of objective performance benchmarks. We suggest instead that a set of principles may be a more flexible way to ensure simulators meet user needs and are fit-for-purpose. These principles should reflect the needs of not just users but audiences, decision makers and other key stakeholders in the Australian fire simulator ecosystem.

Defining and benchmarking fire simulator criteria

Distinguishing simulator user requirements according to a set of key criteria was a valuable approach for users to articulate their specific needs. This project gathered evidence of the importance of a range of different criteria (Table 9). Project participants were supportive of the development of benchmarks for these criteria, but articulating these benchmarks proved to be challenging. When asked for clear and objective standards to guide simulator evaluation or development, many users responded by saying 'it depends', harking back to the complexity of the environment in which they operate.

TABLE 9 - KEY CRITERIA FOR GOOD SIMULATORS ACCORDING TO USERS

CRITERIA	DESCRIPTION
Ease of use	Intuitive interface; Reliable under diverse conditions
Speed	Quick fire behaviour simulation, quick inputs, quick outputs
Configurability	Users can easily adjust models, inputs and outputs
Versatility	Simulator handles all aspects of fire behaviour (Some prefer specialists)
Robustness of modelling framework	Easily handles ensembles and uncertainty; Results are traceable
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
Scale	Appropriate spatial and temporal resolution
Validation	Simulators and underlying models are accurate and validated
Support	Training, documentation and support to run simulator and interpret output
Trustworthiness	Users and agencies trust the simulator, developers and underlying research
Compatibility	Simulator is compatible with other models, agency systems and policies
Value for money	Simulator investments provide value for money

Can the key criteria be benchmarked?

Participants typically found it easiest to attach stand-alone benchmarks to specific performance-based criteria. Four performance-based simulator criteria were identified from participant discussions as possible options to benchmark. These were fire characterization accuracy, speed, resolution and ensembles. Median thresholds for each of the four types of fire characterisation accuracy considered (rate of spread, area burned, fire behaviour and impacts) consistently sat between 60-90% (Table 7). Whilst such benchmarks may be valuable for broadly articulating average user requirements, they are limited in that they (a) inherently overlook the requirements of some users by focusing on the median use (b) risk overlooking the complexity of the many contextual factors that influence the accuracy of a simulator beyond the simulator itself, such as user skill or the quality of the data (c) do not necessarily reflect realistic expectations of output accuracy. Further work is required if any of these criteria are to be operationalised in quantitative terms.

For the remaining three performance criteria - speed, resolution (time-step and spatial resolution) and ensembles – responses varied widely across users. This meant that creating thresholds for ideal use was impracticable (Table 8). Participants were unable to agree upon consistent, measurable indicators for the remaining criteria. These included set-up, technical outputs, validation, operational support, interpretation support, reputation and compatibility. Often these criteria were discussed in terms of the presence or absence, or the increase or decrease, of specific features or capabilities with respect to their current performance. Contextual factors at various levels (e.g. role, organisation, jurisdiction, nation) were difficult to quantify but cited as highly influential in determining people's expectations of the different criteria. The principles and priorities for simulator development outlined later in this report put forward a range of suggestions for how to engage with these criteria in the absence of clear benchmarks.

Categorising fire simulator users

Understanding who uses fire simulators may help to inform simulator development by grouping participants according to consistent ways that they engage with the tools. The two most relevant user categorisations that emerged were use case and jurisdiction (which relates to organisation and geography). Both factors were shown to have implications for how participants used simulators, their requirements from simulators and their current capacities to make the most of these tools.

Use case

Whilst the list of key criteria (Table 9) reflects the factors that were widely agreed to be important for good fire simulators, use case was, at times, a useful indicator of the particular importance of certain criteria for participants. These patterns of importance reflected the types of work that each of the use cases related to as outlined in each of the four criteria descriptions, which can be found in this report's introduction. Variability according to use case, however, was only evident qualitatively, in participant discussions. When participant perspectives were quantitatively measured in the questionnaire, none of the criteria showed clear benchmarks according to use case (Figure 1 and Figure 3). The diversity and complexity within user groups, along with participants consistent use of simulators for tasks outside of their primary use case also meant that the differences identified were generally indicative rather than definitive. Use across multiple use cases is particularly true for fire agencies, who house a significant proportion of Tactical and Strategic operators who can employ simulators in a wide variety of ways outside of their primary use case. These results suggest that use case provides a valuable indication of what their simulator needs might include, however it does not necessarily reflect *all* their needs and therefore cannot be used as an unequivocal template for simulator design or implementation.

All Users

All users value a simulator with a good reputation based on trust in the underlying research.

Tactical Users

Tactical users value a simulator that is fast, easy to use and well supported

Tactical users are less concerned with the ability to investigate multiple specific aspects of fire behaviour and open source software

Strategic Users

Strategic users value a simulator that is configurable, has a robust modelling framework and generates a range of high quality outputs

Strategic users are less concerned with speed and open source software

Research Users

Research users value a simulator that is configurable, has a robust modelling framework and can represent multiple specific aspects of fire behaviour

Research users are less concerned with speed and compatibility with agency settings

FIGURE 12 – SUMMARY OF USE CASE PRIORITIES

Jurisdiction

Jurisdiction, or geographical location, also proved to be an important point of difference between users. It influenced both logistical and cultural differences in how users engaged with simulators and their requirements from them. Similar to use case, the importance of jurisdiction was only qualitatively evident throughout participant discussions and interviews. The questionnaire showed no clear quantitative indication of the importance of jurisdiction or location. In saying this, jurisdiction also importantly appeared to dictate users' capacity to engage in discussions about simulator development in the first place. Underlining this challenge were the questionnaire results, which skew towards participation from the more densely populated south-eastern Australian jurisdictions. Care should be taken in drawing conclusions about smaller jurisdictions due to smaller sample sizes. Table 10 outlines the many jurisdictional and geographical factors that were said to influence the purpose a simulator is used for, the relevance and value of certain simulator capabilities and the support systems most needed. These differences and, at times, inequalities across jurisdictions pose a significant challenge for developers to effectively meet the diversity of user needs. Understanding not only how simulators should be designed for different jurisdictions, but also how engagement, development and management processes that surround fire simulators must vary will be an important area of consideration for developers and decision-makers going forward.

TABLE 10 - SOME JURISDICTIONAL DIFFERENCES RELEVANT TO FIRE SIMULATOR USE AND DEVELOPMENT

THEME	FACTOR
Climatic and environmental conditions	Frequency and severity of fires Consistency/reliability of weather conditions Consistency of vegetation and topography Fuel loads Vegetation types present in landscape
Capabilities	Number of simulator operators relative to area/number of fires Overall size of ground crews User experience/skill level – rotating roles vs permanent Long-term staff transience Capacities/resources to train and support staff and to support simulator transitions Capacity to engage in simulator development Leverage in simulator development discussions
Social context	Risk level Number of stakeholders involved in a fire response Culture around fire risk Stakeholder trust in simulator outputs
Fire management techniques and strategies	Fuel management plans – frequency, coverage, bounded/unbounded burns Tactics for fire suppression
Regulations	Regulations around fire suppression
Tools	Corresponding software/programs used in collaboration with simulators Access to and quality of satellite imagery Type, amount and quality of data on past fires Type, amount and quality of data on vegetation type, fuel load, weather etc. Internet connectivity

Engaging with the fire simulator ecosystem

It is clear from this research that what makes a simulator 'good' extends well beyond the simulator itself. A good simulator must exist within a functional ecosystem. Fire simulator users operate in a complex environment, integrating a wide range of knowledge and experience into their practice. A huge variety of local, organisational, jurisdictional and national factors and stakeholders beyond simulator users also affect outcomes from simulators. For Tactical and Strategic users, a good ecosystem meant a reliable and adaptable platform that is institutionally supported (e.g., by fire and land management agencies) and has the necessary quality of inputs, training, technical support, ongoing research and user community for them to achieve good outputs and outcomes. Similarly, for researchers, a good simulator was a reliable and adaptable platform that is institutionally supported (e.g., research funding bodies, fire and land management agencies) and supported by ongoing research and an active user community. Across the board, participants supported the need to engage with these other deeply interconnected factors such as resources and financing, inputs and the interests of other stakeholders such as data creators and audiences. Any future consideration for how to improve the outcomes from simulators must work to strengthen the broader fire simulator ecosystem and not just the physical simulating tools.

Establishing a framework for fire simulator development

Categorising users and creating benchmarks or thresholds for simulator development may help to regulate simulator development to ensure that the technology meets a certain minimum level of quality. Two factors, however, limit the value of such information. Firstly, participant expectations on what makes a good simulator were so highly varied that generalising about user needs was extremely challenging. It also inevitable that these requirements will continue to evolve over time. Secondly, the wildfire simulator sector in Australia is small and support for simulators is highly dependent upon reputation and trust. Trust relies on relationship building, connection and clear communication. Consistent and effective engagement between users, audiences, developers and other stakeholders is a critical part of this.

A recommendation of this project is the development of a framework that includes a set of broad guidelines or principles for future simulator development as well as a list of current and future priorities. This could form the basis of a sector-wide vision for simulator development and may be useful for building structure and strategy around future stakeholder engagement. Unlike standards or benchmarks, such a framework should be iterative and responsive to a changing environment, community needs and technology, with a focus on ethics and values. A sector-wide development framework would not replace the need for simulator standards or simulator-specific design frameworks but would instead provide guidance to both creation and ongoing maintenance of simulators. A framework or set of principles for simulator development may help to engage with and communicate the complexity and diversity inherent in fire simulators use in Australia and help to increase the fitness-for-purpose of fire simulators. The principles and priorities for simulator development outlined below indicate several areas that could be included in such a framework.

Principles and priorities for fire simulator development

Four principles for fire simulator development were developed from participant findings (Table 11). Below, each of the principle is discussed in detail along with priorities and actions to consider.

TABLE 11 – PRINCIPLES, PRIORITIES AND SUMMARY OF ACTIONS TO CONSIDER

PRINCIPLE	PRIORITIES	ACTIONS TO CONSIDER
Drive simulator performance through improved modelling and data.	*Improve simulator capability*	• Maintain a pipeline of short- and long-term improvements to models and modules, based on science and user needs.
	Improve the availability and quality of input data	• 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. • Clarify and streamline definitions and units around input data.
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 key points such as 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 validation and verification 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, e.g. end of shift, incident and season reviews. • Ask stakeholders the level of engagement they have capacity for • Reflect upon the logistics and resources needed for ongoing, long-term stakeholder engagement.
	Build capacity	• Develop strategies for capacity building tailored to user needs.

Drive simulator performance capability through improved fire behaviour modelling and improved quality and coverage of input data

Priorities

Improve simulator capability
There are a huge variety of simulator capabilities that must be improved or included in future simulators. These relate to the improvement of underlying fire behaviour models, the model interface with data and the modelling process. The prioritisation of these however, varies greatly between users. These capabilities must constantly be improved upon through frequent engagement with users and decision-makers to understand what they need most. Some capabilities relate to the front-end-user interface whilst others are related to the back-end modelling.

Participant suggestions

- Integration of live data in a simple way - dynamic connective feedback of live data - automatically pulls in best/newest knowledge spotting processes, meteorological information and effects (e.g. Coupled atmospheric fire modelling), satellite hotspot data in a suitable timeframe.
- Customisable fuel model – Simpler way to alter fuel loads, arrangement and availability.
- Improved ensemble capacity - Capacity to do rapid tests of ideas and compare multiple different strategies for real-time tactical use.
    - Flexibility to tweak and customise models according to operational importance – e.g to focus in on a particular asset and customise possible results ahead of a high fire day, such as on the fly estimates of the outcomes at a certain wind speed.
- Capacity to integrate manual fire behaviour prediction methods into the same tool
    - Ellipse method used nationally for manual fire behaviour prediction – have automated ellipse construction and adjustability with all the same data layers that are used for the simulator, visible and presented in the same output format.
    - Enabling fire prediction on top of fire line scan data
- Improved model reconstruction capability - A pipeline of high-quality reconstructed fires against which to evaluate simulators to kickstart model development
- Improved capacity for developing and testing simulators for uses at different spatial-temporal scales
    - finer scales to simulate low-intensity fires such as prescribed burns, back burns,
    - extreme fire behaviour
- Improved modelling capabilities
    - Modelling of growth of fires from a point, speeds of fire growth – all simulators mostly using quasi-steady rate of spread of the fire – wind direction and changes of wind direction
    - Incorporation of smoke modelling
    - Better modelling of all fuel types (e.g. pine plantations, mallee heath)
    - Spotting model included
    - Suppression to be built into modelling

Please note: Whilst this list can provide guidance, it is by no means exhaustive. Users had a huge number of suggestions for simulator features and ideas about the future of simulator development. It is important to note participants were acutely aware of the financial and bureaucratic constraints surrounding the prioritisation of new simulator features (as such, these ideas were typically put forward as requests rather than demands).

Actions to consider

Engage frequently with users, decision-makers and audiences to understand where development improvements are required or may be beneficial.

Be open to advancements in technology and scientific understanding to improve both front-end-user interfaces and back-end modelling components.

Assess the benefit to cost ratio of changes to simulators, in particular, the time-commitment required of users to take on the updates.

Improve the availability and quality of input data

Improved data is hugely important for simulator users and is central to improving outcomes and trust in simulator outputs. The need for increased availability and quality of input data was frequently mentioned as a limitation of existing simulator systems. Jurisdictional/regional differences were also often mentioned as influencing the quality and availability of data.

Participant suggestions for improved data quality

- Improvements to data quality are particularly necessary for fuel data, such as national fuel mapping of load and structure and in remote locations.
- Support development of dynamic real-time data about suppression such as breaks, backburning and Large Aerial Tanker lines.
- Support future capabilities for the automated collection of data following events and simulations, which will facilitate learning and refining knowledge of fire behaviour.
- Ensure the improvement of data for validation, including the improved collection of data on case studies from past fires.
- Ensure consideration of appropriateness of scale of data.

Participant suggestions for improved data storage and access

- Centralised data sources - Access across agencies and organisations.
- Reducing the amount of data required to feed into a program is important. Improved ability to compress data or extract only necessary or relevant data will improve simulator useability. This is particularly important for regional and remote users.
- Better and easier ways to identify important data and information.

Actions to consider

Expand existing data collection methods to provide better information on inputs, for example, fuel mapping, real-time suppression activities and case-study data of past fires.

Improve data networks to ensure greater access, availability and useability of input data.

Address and remove regional and jurisdictional differences in data access and quality.

Build automated data collection into simulators for user-support and ongoing refinement of simulating tools.

Minimise the impact of inappropriate scale and resolution for specific simulation purposes and contexts by, for example, improving user knowledge of appropriate data use and greater standardisation of simulator inputs.

Improve usability of fire simulator software and hardware, including platform stability, outputs that facilitate effective communication with audiences and ongoing support for users and partners.

Priorities

Ensure user-centric design in development of simulator interfaces.

Along with improving the capabilities of fire simulators, it is important to improve the functionality and useability of these tools. This relates to the Graphical User Interface and the command line interface option, for power user/high performance computing. Debate surrounds whether a system should prioritise configurability or automation as well as the expert or general user. Whilst it may be ideal to have a large suite of simulator programs available for different uses each suitable for different types of users, a fire agency for example, will likely focus their limited resources on one or two simulators, even though their workforce encompasses a wide variety of uses and skill levels. Simulators will therefore need to be appropriate for a wide swathe of users. This means that if simulator developers hope to meet the diverse needs of users, they will need to account for a diversity of factors and create a system that can be easily tailored. Developers must strike the balance between suitable configurability (easy to edit/toggle functions/sub models) without being too configurable (complex/not enough automated functions). High configurability can be useful for expert or daily users but less so for others. According to questionnaire participants, people used simulators on average only for approximately 20% of their role with some as low as 1%. Users typically want an interface that is intuitive, quick to use and easy to edit.

A simulator's flexibility to incorporate new science, data and innovations is also essential as inputs and technology develop and evolve rapidly. This capacity to incorporate new science should ideally have minimal interference on the front-end-user interface and system functioning. Consistency across versions or systems is integral to make transitions easier for users. Systems that maintain the same user interface whilst conducting incremental changes or fixing problems to the back end are beneficial for a number of reasons: they reduce the training load required of staff; allow users to maintain unofficial tricks, skills and add-ons from previous versions; and sustain user trust and support in the tool.

Actions to consider

Develop intuitive and efficient workflows that are easy to learn.

Built-in support with troubleshooting such as errors clearly stated or prompts may be valuable, particularly for beginners. Mechanisms to provide feedback may also be valuable.

Develop dual modes of 'general' and 'expert' capabilities to meeting diverse user needs. Whilst the capacity to switch between these modes or turn certain features on and off may be ideal, there are likely both financial costs and impacts on computing power associated with such abilities.

Maximise flexibility and adaptability whilst ensuring consistency across versions and updates. The capacity for ongoing and incremental change to incorporate new scientific advancements, data sources and technological innovations is essential. It must occur, however, with minimal disruption to front-end-user experiences. This capacity should ideally be built into a new system from the start.

Improve stability and usability of hardware and infrastructure

Platform stability was consistently identified as significant for participants. The stability of simulator software does not occur in a vacuum. It is significantly dependent on the encompassing hardware as well as the broader infrastructure that supports simulator use. The physical infrastructure must be both consistently stable and suited to the specific needs and capacities of users. Relevant factors include access to high-quality reception or internet for the transfer of large amounts of data, the necessity of mobile or offline capabilities for working in remote areas that are reception 'dead-zones' and the resourcing capacity to support expensive Cloud services and data storage. This is particularly pertinent for remote jurisdictions where the communication of information and the transfer of data can be challenging and costly.

Actions to consider

Communicate frequently with end-users to understand their evolving infrastructure requirements.

Pay particular attention to smaller or remote jurisdictions who may face greater challenges with data communication and system stability. Greater investment in infrastructure may be required in these contexts.

Provide suitable offline alternatives for simulators, when necessary.

Reduce the amount of data, memory and computing power needed to run simulators.

Improve interpretability and communicability of outputs

Simulations are important as both predictors of fire behaviour and communication tools. The way that simulations are communicated and understood is essential. If a simulator output fails to tell a story, then it is not fit-for-purpose. In general, outputs need to be clear, easy to understand and transparent. The communication of uncertainty and assumptions was extremely important to participants and is essential for building trust between the developers, the users and the audience. This additional information also helps to build reasonable expectations of the results and means that the responsibility of deciding the most relevant information is no longer wholly the responsibility of the simulator creator. Poor interpretation and communication of outputs can not only obscure the potential usefulness of simulations but can also risk decreased confidence/trust in simulating tools more broadly.

Actions to consider

Ensure simulators have a range of different outputs that can be tailored for the variety of contexts and end-user requirements.

Ensure simulators clearly articulate why a certain result has been achieved to determine whether the result is credible or if modifications are needed.

Clear communication of supplementary briefing points such as model assumptions and uncertainty.

Develop a standardised reporting format for outputs.

Provide training and support for the end-user to improve their understanding and trust of simulator outputs.

Provide comprehensive support for simulator users and their audiences

Good support systems help users to make the most of a simulator. Good support is essential for transitioning between simulators, particularly when informal user networks are yet to be established. Good support is also important for bridging the skills gap between users, particularly for new staff or those with rotating responsibilities who might rarely use the tools. Future considerations about training should ask whether there is value in defining clear levels of competency, what those levels should be and if they should be accredited or nationalised.

Actions to consider

Consider implementing a variety of support options for users and audiences such as an in-depth training guide or manual, a support hotline, an easy to understand troubleshooting support built into the GUI, a digital troubleshooting forum, an interactive workshops/online forums, an in-house mentoring and support videos.

Agree upon whether training should be nationalised or accredited.

Adopt a comprehensive and transparent approach to the validation and verification of simulators and their inputs.

Priorities

Establish validation and verification standards

The wildfire and emergency management sector would benefit from a clear and systematic approach to validation and verification of inputs, models and simulators. This approach should be broadly accepted by the simulator community and must ensure that only good science and high-quality data is accepted. The question of what is 'good science' however, is deeply complex and whilst peer-reviewed local and international literature was seen as valuable, so too was grey literature, in-house documentation, operational and lived experience of users and audiences. Considering the roles of each of these types of knowledge will be challenging but important. The need for benchmark testing against historical fires is also necessary. Creating these standards is essential for the systematic assessment of new simulators and their inputs as well as for building trust in outputs.

These standards should also do the following:

- Outline under what circumstances validation needs to happen.
- Provide clarity around if and under what circumstances, a simulator user has flexibility to tweak inputs without the need for validation/peer review.
- Articulate when simulator features or inputs are based on peer-reviewed science/data and when they are not.

Actions to consider

Developers, users and decision-makers must work together to establish standards for evaluation, verification and validation of simulators, models and data.

Developers, users and decision-makers must also establish a mechanism for reviewing and updating these standards as the technology and contexts of simulator use evolve over time.

Emphasise transparency and traceability

Transparency and traceability are linked to communication and are also important for building reputation and trust. Documentation is a valuable tool for the systematic communication of information between developers, users and audiences. It can provide rigour after an incident, it can help people to learn and understand why decisions were made, it improves accountability and provides opportunities for feedback to improvement. Participant discussions on documentation were divided into three areas,

- Simulator function – This includes the process by which a simulator was developed and validated and the science and knowledge behind the system. It also includes general information about how a simulator works and how to use it.
- Simulator use process – This encompasses information about how a simulator was used during the generation of a singular output. It includes input data, decisions, assumptions and details of outputs. This documentation addresses user concerns about simulators being a black-box by providing information about why, in the given context, a simulator acted in the way that it did. Users are interested in being able to trace key drivers of outputs. It also includes information about how long it took to create and run, plus if there were issues or errors. Often, a simulator user will need to manually collate this data, which can be time consuming. In-built collection of this data instead could occur automatically as part of the simulation process.
- Simulator outcomes – This documentation extends beyond the simulator itself to include audiences and decision-makers to understand how a simulator output affects decision-making and outcomes. It requires reflection and review following an incident to determine the effectiveness of a simulation. Such information would be valuable for case study analysis alongside the simulator 'use process' data.

Actions to consider

Establish guidelines or expectations for creating the three types of documentation - simulator function, simulator use processes and simulator outcomes.

Integrate the automatic collection of 'use process' data alongside the running of a simulation.

Reflect on how best to integrate outcomes from simulator use into data collected about the simulator use process. This could be, for example, a feature that allows users to easily input information about simulator outcomes post-event.

Focus on outcome-oriented evaluation and development

The overall quality of a simulator should be ultimately be judged on how it is assists and influences decision-making and the outcomes of decision-making. Whilst it is important to evaluate the technical aspects or model/simulator performance, changes or 'improvements' are of limited value unless the outcomes they lead to are also better. Several participants stressed that, in some circumstances, focusing solely on improving outputs can make outcomes worse.

The effect of a particular simulator on outputs is not always the same. It can vary between contexts and can evolve over time. The effect on outputs is also caused by a variety of different factors that can either relate directly to the simulator itself or can be linked instead to the broader simulator ecosystem. It is important to create mechanisms to evaluate where exactly improvements need to be made. For example, immediately following a change to a model or simulator, there may be a decrease in outcome quality. This may be due to problems with the tool itself, but it may also be that users are not yet properly trained on the new simulator or feature.

Assessing the outcomes of a simulation could possibly be built into the system itself by asking

- The simulator user/operator to provide feedback on whether the tool successfully supported and enhanced their responsibilities.
- The end-user (audience or client) to provide feedback on,
    - Whether they experienced benefits from the output (e.g. Did they consider the information provided to them? Was it valuable? And did they make a better decision because of it?)
    - Whether the event circumstances changed on the ground (e.g. What improved, worsened, or was not affected?).

Actions to consider

Prioritise good outcomes over good outputs as the primary driver of simulator development decisions.

Engage with the simulator stakeholder community to understand and document the effect of simulator outputs on outcomes. Reflect on how the findings can be used to inform simulator development decisions and consider whether the collection of this information can be made systematic.

Develop a simulator feature that allows users to easily input information about simulator outcomes post-event.

Maintain a cohesive approach to development and use through effective governance, capacity building and inclusive and ongoing consultation between users, audiences and developers.

Priorities

Establish effective governance and a long-term vision for fire simulator use and development

The wildfire research and management sector would benefit from a well-coordinated and collaborative long-term vision for future fire simulators to allow for continuous and proactive development and implementation that causes minimal disruption for users. This will require significant financial investment from government and commitment across the board from the simulator community. A shared, inclusive and well-articulated vision will also help to avoid development simply "reinventing the wheel" (Interview participant), ensuring that improvements made are impactful and relevant.

Governance priorities according to participants

- Establish a clear long-term vision for future research and development
    - Re-alignment between short term priorities and long-term development
    - Identifying current and future research needs
- Understanding how the fire sector fits and works together
- Common research to operational pathway to prevent siloing of research
- A step-by-step strategy of how to integrate a new system as the new science becomes available, as improvements come in.
        - Which outlines what the new system/tools do and how it works better.
- Pathways and guidelines for development and capacity building of users
- Consider where bureaucracy is helpful and when it hinders development – How to ensure quality whilst maintaining flexibility?

- Ensure inclusivity of diverse stakeholders and reduce jurisdictional biases in development.
- Formalising communication channels
    - between simulator users/operators and their end-users/audience to discuss the interpretation and consequences of the outputs.
    - Communication and engagement between stakeholders during simulator development.
    - Nationalised network for simulator use and interpretation support.
    - International links- It can be valuable to stay in touch with other fire prone jurisdictions to find out how their simulators are going.
- Consideration of the following must occur
    - Prioritisation of limited resources- technical vs social focus for system improvements
    - Is it better to have multiple models/simulators?
    - Is nationalisation a good idea (Appendix 3)?
    - Ownership and responsibility over outputs
    - How best to communicate uncertainty

Actions to consider

Establish a clear long-term vision for future research and development of fire simulators, models and data.

Develop strategies for the integration of new science, capacity building, communication and the functioning of the fire simulator network.

Identify critical purposes for simulator use

"I think I would still like to unpack the question of what do we want a fire simulator to do? What do we actually want?... That then is the framing for the question of what makes a good one." Workshop 1 participant

There is value in building greater clarity and agreement across the wildfire sector about the activities that fire simulators are currently most impactful in supporting and will be in the future.

This research determined three major use cases: Tactical, Strategic and Research. Additional use cases include for training and community engagement. Within these three major use cases there are a variety of common uses for simulators (Table 1). Decision making around major investments in fire simulators requires clarity about which, if any, are highest priority and where the greatest benefits are likely to be achieved. Some participants suggested that simulators are often used for purposes beyond what they were initially intended for. Whilst this is not always a bad thing, reflecting on origins and intended purpose of a tool it helps us be aware of whether that tool is appropriate or if alternatives could be better suited to the new purposes.
Strategic future simulator development means that the clarification of these purposes for simulator use must be embedded in the context of rapidly technological improvement and climate change impacts predicted for the future. Clarifying purpose also means aligning definitions relevant to fire simulators – e.g. What do we mean by good 'performance'?

Actions to consider

Continually engage in cross-sector stakeholder discussions to understand current and future purposes for simulators.

Reflect on the origins and intended purpose of a fire simulator to assess whether this current tool is appropriate for certain purposes or if alternatives may be better suited.

Be open to new possible uses for simulators. For example, several participants suggested that fire simulators could be more greatly used as part of wildfire mitigation. Using simulators for this purpose would provide an opportunity to shift attention away from emergency response and towards longer-term planning of wildfire mitigation and prevention.

Identify principal ethics to guide simulator development and use
Many of the priorities discussed throughout this report have ethical dimensions. In the interests of all parties, these dimensions must be proactively and openly considered and addressed.

These ethical questions have important implications for numerous factors related directly to the development of simulators including design, training and support. They also have ramifications for factors around trust and legal accountability. Multiple participants noted that there is always a chance, due to the degree of high risk and consequence of decision-making around wildfire, that simulator users and audiences may need to justify their actions in a court or coronial inquiry.

Actions to consider

Stakeholders across the wildfire community must collectively engage with ethical questions, such as the following:

Access - Who should have access to simulators and their outputs?

Resourcing- Often capacity is as much of a barrier as access. To what extent is the broader simulator community responsible for ensuring that all users have the capacity and resources to make use of simulator access?

Responsibility - Who is responsible and accountable for simulators and their outputs?

Human-Computer interactions - How should we manage human-computer interactions, particularly as technology changes and improves?

Ensure inclusive consultation and collaborative development
Consultation is important for understanding the many ways that people use a tool, including the 'unofficial' ways outside of the tool's intended design. Good consultation, however, is challenging and, when done poorly, can be ineffective or detrimental.

There are a wide number of stakeholders relevant to ensuring that simulators are high quality. These include developers, users, audiences and the scientists responsible for building the models and collecting data. Consultation with these different stakeholder groups would ideally occur across all stages of simulator development and upgrades. Affordances must be made to actively engage minority users, such as remote users or smaller jurisdictions. Often these users do not have capacity to participate in consultation processes and their voices go unaccounted for. This is important for building reputation and trust and for ensuring that the tool is suitable for its intended purpose.
Failure to properly consult with users may mean that users continue to use an old version due to lack of suitability of or trust in the new option. When introducing changes, it is also important to communicate these changes such as, gain/loss of functionality or complexity; whether the science behind the change been peer-reviewed and published; and how and for what purposes, simulators and subsequent versions have been developed.
Whilst insufficient engagement was cited by several participants as limiting the potential for fire simulators to meet operational needs and expectations, there were exceptions to this view. Some users suggested that, in particular circumstances, broadscale engagement was, in fact, detrimental. It was suggested instead that the particular needs of a small subset of users can be more greatly met through targeted streamlined development, particularly when tied with support availability for users.

Whilst it is important that simulator development teams take responsibility for collaboration and engagement, it also requires a longer-term sector-wide response. Consideration must be given to whether there should be ongoing formal, systematic process for collaboration and engagement. If this is deemed to be important, this raises questions about whether developers along with other stakeholders, are suitably resourced for such engagement and what suitable resourcing might look like.

Actions to consider

Actively seek and incorporate input from a wide range of stakeholders. Appendix 4 provides additional information and lines of questioning to support understanding of stakeholder context when considering simulator change.

Reflect upon the logistics and resources needed for ongoing, long-term stakeholder consultation. This pertains to both the sector-wide planning for simulator development and small-scale changes to specific simulators. How can simulator development and use effectively engage with relevant stakeholders given the practical and financial challenges and limitations that often come with thorough or large-scale consultation?

Build capacity

Staff capacity building is important for integrating a simulator into business-as-usual operations. This requires improving trust and support for the tools and increasing staff skill, knowledge and experience with the tools. Training and support must be developed in collaboration with the simulator development team and, importantly, engagement with these services must be encouraged and facilitated by the organisations using these programs. Inappropriate use of simulators and user error, typically a result of inexperienced use, can have a significant effect on broader trust of simulators and their outputs. Some participants suggested the need for differentiated training for both different types of users and audiences.

Education will be beneficial for the following reasons,

- To improve stakeholder understanding of the limits, uncertainty, assumptions and limitations of fire simulators.
- To ensure simulators are not used inappropriately in situations they were not designed for and are not suited to.
- To help stakeholders understand and communicate that a simulator is not a perfect representation of reality. “Too many people think that the simulation output is a reality and so instead of going to look at what the reality is, they just take the simulation as the reality” (Interview participant).

Actions to consider

Reach sector-wide agreement about pathways and guidelines for stakeholder capacity building that recognise the diverse needs and capacities of users. This can sit as part of a broader governance strategy for future simulator development and use.

Make sure that simulator-specific support offerings reflect sector-wide strategies for capacity building.

Limitations

This engagement project generated a large amount of rich feedback from users about their priorities and needs. We thank them for their time and expertise. Nevertheless, it is important to recognise that knowledge gaps remain in our understanding of fire simulator development and use. Some of these relate to the initial project EOI, for instance a comprehensive catalogue of the decisions that simulator users make when using these tools. Other gaps related to user groups we were unable to engage in-depth, such as the audiences of simulator users and the agency staff responsible for making major investment decisions about simulator resourcing. Other gaps were identified during the project and represent promising avenues for future inquiry, such as the potential to regularly track the outcomes of simulator use. Ultimately, this project demonstrated that in order to meet the diverse, context-dependent and evolving needs of fire simulator users, we need principles, tools and processes that are correspondingly nuanced and adaptable. Importantly, this does not mean that major new investments in user engagement or capacity building are required. Many of the priorities identified here can be pursued through existing processes and light touch mechanisms.

The work should be interpreted in the context of other research in this area, including ongoing agency evaluations of Spark and other fire simulators in multiple states and territories, the culture of simulator use (e.g. Neale et al. 2021) and Natural Hazards Research Australia projects on Predictions in Public (about the design, communication and dissemination of predictive maps to the public), uncertainty in bushfire spread prediction, enhanced decision making in emergency management and engagement to enhance severe weather impact predictions.

Future directions

Areas of potential investigation related to fire simulators and the fire simulator ecosystem include:

- Testing success indicators and evaluation frameworks developed for other model types identified in the literature. Assess how effectively they work on fire simulators to inform development and investment priorities.
- Better understanding the particular needs and priorities of smaller jurisdictions and stakeholders with limited capacity to participate in simulator design and development.
- Investigating the outcomes of simulator use e.g. decisions, events (i.e. evaluating models according to fitness for purpose rather than their ability to mirror reality).

Areas of potential investigation related to specific stakeholders and issues include:

- Quantitative analysis of agency response time e.g. using agency metadata on portals like FireWeb to explore time taken to produce outputs from initial ignition, run speed and to better understand what factors influence the effectiveness of outputs.
- Detailed qualitative analysis of how Tactical users (FBAns) use simulators, skills required now and, in the future, how they can be supported and what affects how FBAns engage with simulators (e.g. ‘good’ vs ‘bad’ fire days, busy seasons, fire size, access to information, challenging weather conditions).
- Identification of potential new users and uses of fire simulators.
- Understanding decision making processes including the influence of cost, the changing cost profile (c.f. computational storage and processing).
- Understanding simulator design constraints, emerging and cutting-edge options.
- Understanding different audience types and their information needs, including links to the public (c.f. Predictions in public project).
- SWOT analysis of current and emerging input datasets, governance requirements, the role of peer-review in dataset creation.
- The potential role of artificial intelligence, big tech companies and automation, including implications for hazard prediction, user acceptance and communication.
- Collaborative governance and ethical issues in the control of, access to and trust of simulators.

Project team

The research project was coordinated by researchers at the University of Melbourne and Deakin University.

Research team

Caitlin Symon, University of Melbourne
Hamish Clarke, University of Melbourne
Timothy Neale, Deakin University
Gabrielle Miller, Deakin University
Kate Parkins, University of Melbourne
Erica Marshall, University of Melbourne
Alex Filkov, University of Melbourne
Trent Penman, University of Melbourne

End-users

Thomas Duff, Victorian Country Fire Authority
David Field, NSW Rural Fire Service
Simon Heemstra, NSW Rural Fire Service
John Bally, AFAC

Natural Hazards Research Australia

George Goddard, Natural Hazards Research Australia
Kat Haynes, Natural Hazards Research Australia

References

Bakhshaii A & Johnson EA (2019) A review of a new generation of wildfire–atmosphere modeling. Canadian Journal of Forest Research, 49(6), 565-574. https://doi.org/10.1139/cjfr-2018-0138

Beven K & Lane S (2022) On (in) validating environmental models. 1. Principles for formulating a Turing-like test for determining when a model is fit-for purpose. Hydrological Processes, 36(10), e14704. https://doi.org/10.1002/hyp.1470

Bowman D et al. (2017). Human exposure and sensitivity to globally extreme wildfire events. Nat Ecol Evol 1, 0058 (2017). https://doi.org/10.1038/s41559-016-0058

Cortes CAT et al (2023) 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. 30(7):4285-4303. doi: 10.1109/TVCG.2023.3258440

Cruz MG et al. (2014) Fire behaviour knowledge in Australia: a synthesis of disciplinary and stakeholder knowledge on fire spread prediction capability and application. Bushfire Cooperative Research Centre. https://www.bushfirecrc.com/sites/default/files/managed/resource/fire_knowledge_synthesis_final_report.pdf

Cruz MG et al (2018) Assessing improvements in models used to operationally predict wildland fire rate of spread. Environmental Modelling and Software, 105, 54-63 https://doi.org/10.1016/j.envsoft.2018.03.027

Cunningham, C.X., Williamson, G.J. & Bowman, D.M.J.S. Increasing frequency and intensity of the most extreme wildfires on Earth. *Nat Ecol Evol* **8**, 1420–1425 (2024). https://doi.org/10.1038/s41559-024-02452-2

Di Virgilio, G et al. (2019). Climate change increases the potential for extreme wildfires. Geophysical Research Letters, 46, 8517–8526. https://doi.org/10.1029/2019GL083699

Doerr SH & Santín C (2016) Global trends in wildfire and its impacts: perceptions versus realities in a changing world. Phil. Trans. R. Soc. B 371(1696). http://dx.doi.org/10.1098/rstb.2015.0345

Duff et al. (2016) Indices for the evaluation of wildfire spread simulations using contemporaneous predictions and observations of burnt area. Environmental Modelling and Software, 83, 278-285 https://doi.org/10.1016/j.envsoft.2016.05.005

Duff et al. (2018) Conditional performance evaluation: using wildfire observations for systematic fire simulator development. Forests, 9(4), 189; https://doi.org/10.3390/f9040189

Duff TJ et al. (2018) Conditional Performance Evaluation: Using Wildfire Observations for Systematic Fire Simulator Development. Forests, 9(4):189 https://doi.org/10.3390/f9040189

Faggian et al. (2017) Final Report: An evaluation of fire spread simulators used in Australia. Bushfire Predictive Services. 1.18
http://www.bom.gov.au/research/publications/otherreports/FPS_Final_Report_v1.81_Evaluation_Of_Simulators_Release.pdf

Filippi et al. (2014) Representation and evaluation of wildfire propagation simulations. International Journal of Wildland Fire, 23, 46–57
http://dx.doi.org/10.1071/WF12202

Filkov AI et al. (2020) Impact of Australia's catastrophic 2019/20 bushfire season on communities and environment. Retrospective analysis and current trends. Journal of safety science and resilience, 1(1), 44-56. https://doi.org/10.1016/j.jnlssr.2020.06.009

Fox-Hughes P et al. (2024) An evaluation of wildland fire simulators used operationally in Australia. International Journal of Wildland Fire, 33(4), https://doi.org/10.1071/WF23028

Hamilton SH et al. (2019) A framework for characterising and evaluating the effectiveness of environmental modelling. Environmental Modelling & Software, 118, 83-98 https://doi.org/10.1016/j.envsoft.2019.04.008.

Hamilton SH et al. (2022) Fit-for-purpose environmental modelling: Targeting the intersection of usability, reliability and feasibility. Environmental Modelling & Software, 148, doi: https://doi.org/10.1016/j.envsoft.2021.105278

Humber et al. (2020) Assessing the shape accuracy of coarse resolution burned area identifications. IEEE Transactions on Geoscience and Remote Sensing, 58(3), 1516 - 1526 https://doi.org/10.1109/TGRS.2019.2943901

Katan and Perez (2021), ABWiSE v1.0: toward an agent-based approach to simulating wildfire spread. Natural Hazards and Earth System Sciences. 21, 3141–3160 https://doi.org/10.5194/nhess-21-3141-2021

Kelly LT et al. (2023) Understanding fire regimes for a better Anthropocene. Annual Review of Environment and Resources, 48(1), 207-235 https://doi.org/10.1146/annurev-environ-120220-055357

Kelso JK et al. (2015) Techniques for evaluating wildfire simulators via the simulation of historical fires using the Australis simulator. International Journal of Wildland Fire, 24, 784–797. doi:10.1071/WF14047

Matthews KB et al. (2011) Raising the bar? The challenges of evaluating the outcomes of environmental modelling and software. Environmental Modelling and Software, 26(3), 247-257 https://doi.org/10.1016/j.envsoft.2010.03.031

McIntosh BS et al. (2011) Environmental decision support systems (EDSS) development – Challenges and best practices. Environmental Modelling and Software, 26(12):1389–1402, http://dx.doi.org/10.1016/j.envsoft.2011.09.009

Merritt WS et al. (2017) Realizing modelling outcomes: A synthesis of success factors and their use in a retrospective analysis of 15 Australian water resource projects. Environmental Modelling and Software, 94, 63-72 https://doi.org/10.1016/j.envsoft.2017.03.021

Morss RE et al. (2005) Flood risk, uncertainty and scientific information for decision making: lessons from an interdisciplinary project. Bulletin of the American Meteorological Society 86 (11): 1593-1601. https://doi.org/10.1175/BAMS-86-11-1593

Morvan et al. (2022) Fifty years of progress in wildland fire modelling: from empirical to fully physical CFD models, 350, 1-9 http://dx.doi.org/10.5802/crmeca.133

Neale T & May D (2020) Fuzzy boundaries: expertise and culture in bushfire prediction. Social Studies of Science 50:837-859. doi: 10.1177/0306312720906869.

Neale Timothy, Vergani M, Begg C, Kilinc M, Wouters M, & Harris S (2021) 'Any prediction is better than none'? A study of the perceptions of fire behaviour analysis users in Australia. International Journal of Wildland Fire, 30(12), 946-953. doi: https://doi.org/10.1071/WF21100.

Neale T (2023) Performing control: Ritual and divination in Australian emergency management. American Ethnologist, 50(4), 568-581. doi: 10.1111/amet.13218.

Ozaki, M, Harris R M, Love P T, Aryal J, Fox-Hughes P and Williamson GJ (2022). Impact of vertical atmospheric structure on an atypical fire in a mountain valley. Fire, 5(4), 104-122. https://doi.org/ 10.3390/fire5040104

Pacheco AP et al. (2015) Cohesive fire management within an uncertain environment: a review of risk handling and decision support systems. Forest Ecology and Management, 347, 1-17. https://doi.org/10.1016/j.foreco.2015.02.033

Pais S et al. (2023) The REMAINS R Package: Paving the way for fire-landscape modelling and management. Environmental Modelling and Software, 168, https://doi.org/10.1016/j.envsoft.2023.105801

Parker WS (2020) Model Evaluation: An Adequacy-for-Purpose View. Philosophy of Science, 87(3), 457–477. doi:10.1086/708691

Peace et al. (2021) BNHCRC Couple d fire-atmosphere simulations of five Black Summer fires using the ACCESS-Fire model. BNHCRC report. Report no. 705.2021 https://www.bnhcrc.com.au/sites/default/files/managed/downloads/coupled_fire-atmosphere_simulations_black_summer_final_report.pdf
Penman TD et al. (2020) Effect of weather forecast errors on fire growth model projections. International Journal of Wildland Fire 29, 983-994. https://doi.org/10.1071/WF19199
Penman TD et al. (2022) Improved accuracy of wildfire simulations using fuel hazard estimates based on environmental data. Journal of Environmental Management. 301, https://doi.org/10.1016/j.jenvman.2021.113789.
Plucinski, MP et al. (2017) Improving the reliability and utility of operational bushfire behaviour predictions in Australian vegetation. Environmental Modelling & Software, 91, 1-12. https://doi.org/10.1016/j.envsoft.2017.01.019
Prichard SJ et al. (2023) The REBURN model: simulating system-level forest succession and wildfire dynamics. Fire Ecology, 19(1), 1-32. https://doi.org/10.1186/s42408-023-00190-7
Rapp C (2020) Wildfire decision support tools: an exploratory study of use in the United States. International Journal of Wildland Fire 29(7), 581-594. http://www.publish.csiro.au/paper/WF16038.
Rayner S (2005) Weather forecasts are for wimps: why water resource managers do not use climate forecasts. Climatic Change 69(2), 197-227. https://doi.org/10.1007/s10584-005-3148-z
Sa et al. (2017) Evaluating fire growth simulations using satellite active data. Remote Sensing of Environment, 190, 302-317. https://doi.org/10.1016/j.rse.2016.12.023
Salis et al. (2016) Predicting wildfire spread and behaviour in Mediterranean landscapes. International Journal of Wildland Fire, 25(10), 1015-1032. https://doi.org/10.1071/WF15081
Schumaker NH, Watkins SM & Heinrichs JA (2022) HexFire: A Flexible and Accessible Wildfire Simulator. Land (Basel), 11(8), 1-16. doi: 10.3390/land11081288
Sharples JJ, Cary GJ, Fox-Hughes P et al. (2016) Natural hazards in Australia: extreme bushfire. Climatic Change, 139, 85–99 https://doi.org/10.1007/s10584-016-1811-1
Sun et al. (2023) Facing the Wildfire Spread Risk Challenge: Where Are We Now and Where Are We Going? Fire. 6(6), 228. https://doi.org/10.3390/fire6060228
Ujjwal et al. (2022) A probability-based risk metric for operational wildfire risk management. Environmental Modelling and Software, 148, https://doi.org/10.1016/j.envsoft.2021.105286.
Voinov A and Bousquet F (2010) Modelling with stakeholders. Environmental Modelling and Software, 25(11), 1268-1281, https://doi.org/10.1016/j.envsoft.2010.03.007.
Voinov A & Gaddis EJB (2008) Lessons for successful participatory watershed modeling: A perspective from modeling practitioners. Ecological Modelling, 216(2), 197-207. https://doi.org/10.1016/j.ecolmodel.2008.03.010.
Walling E & Vaneeckhaute C (2020) Developing successful environmental decision support systems: Challenges and best practices. Journal of Environmental Management, 264. https://doi.org/10.1016/j.jenvman.2020.110513.
Wegrzynski W & Lipecki T (2018) Wind and fire coupled modelling – Part I: Literature review. Fire Technology, 54, 1405–1442 https://doi.org/10.1007/s10694-018-0748-5
Wu R et al. (2023) vFirelib: A GPU-based fire simulation and visualization tool, SoftwareX, 23, https://doi.org/10.1016/j.softx.2023.101411.
Zacharakis & Tzihrintzis 2023. Integrated wildfire danger models and factors: A review. Science of the Total Environment, 899, https://doi.org/10.1016/j.scitotenv.2023.165704.

Appendix 1. Performance metrics by use group

TABLE 12 - BENCHMARKS FOR PERFORMACE CRITERIA - FIRE CHARACTERISATION ACCURACY, TIMESTEP, SPATIAL RESOLUTION, OUTPUT SPEED AND ENSEMBLES SHOWING MEAN, MEDIAN AND RANGE FOR TACTICAL USER RESPONSES

VARIABLE	MEAN	MEDIAN	RANGE
Rate of spread accuracy (%) - min	61	60	10-95
Rate of spread accuracy (%) - max	87	90	50-100
Area burnt accuracy (%) - min	58	60	10-90
Area burnt accuracy (%) - max	82	80	40-100
Fire behaviour accuracy (%) - min	58	60	20-90
Fire behaviour accuracy (%) - max	83	80	30-100
Impacts accuracy (%) - min	58	60	10-90
Impacts accuracy (%) - max	84	84	60-100
Spatial resolution (m^2) - min	148	30	1-3000
Spatial resolution (m^2) - max	996	220	30-10000
Model timestep (minutes) - min	29	30	0.02-180
Model timestep (minutes) - max	209	60	10-1440
Output speed minutes) - min	4	1	0-30
Output speed (minutes) - max	64	15	1-1440
Number of ensembles - min	7	1	0-50
Number of ensembles - max	3128	50	3-1.00E+05

TABLE 13 - BENCHMARKS FOR PERFORMACE CRITERIA - FIRE CHARACTERISATION ACCURACY, TIMESTEP, SPATIAL RESOLUTION, OUTPUT SPEED AND ENSEMBLES SHOWING MEAN, MEDIAN AND RANGE FOR SRTATEGIC USER RESPONSES

VARIABLE	MEAN	MEDIAN	RANGE
Rate of spread accuracy (%) - min	66	63	50-85
Rate of spread accuracy (%) - max	88	90	50-100
Area burnt accuracy (%) - min	63	61	40-85
Area burnt accuracy (%) - max	87	90	50-100
Fire behaviour accuracy (%) - min	63	61	40-80
Fire behaviour accuracy (%) - max	87	90	50-100
Impacts accuracy (%) - min	67	70	50-80
Impacts accuracy (%) - max	87	90	50-100
Spatial resolution (m^2) - min	95	30	10-1000
Spatial resolution (m^2) - max	1209	180	30-10000
Model timestep (minutes) - min	14	10	0.5-30
Model timestep (minutes) - max	73	60	1-240
Output speed minutes) - min	6	1	0.02-60
Output speed (minutes) - max	1427	13	0.5-20160
Number of ensembles - min	1169	1.5	1-20000
Number of ensembles - max	12559	54	5-1.00E+05

TABLE 14 - BENCHMARKS FOR PERFORMACE CRITERIA - FIRE CHARACTERISATION ACCURACY, TIMESTEP, SPATIAL RESOLUTION, OUTPUT SPEED AND ENSEMBLES SHOWING MEAN, MEDIAN AND RANGE FOR RESEARCH USER RESPONSES

VARIABLE	MEAN	MEDIAN	RANGE
Rate of spread accuracy (%) - min	53	50	30-90
Rate of spread accuracy (%) - max	83	83	60-100
Area burnt accuracy (%) - min	59	56	36-90
Area burnt accuracy (%) - max	89	90	80-100
Fire behaviour accuracy (%) - min	55	60	30-90
Fire behaviour accuracy (%) - max	85	86	70-100
Impacts accuracy (%) - min	42	40	20-60
Impacts accuracy (%) - max	74	70	49-100
Spatial resolution (m^2) - min	146	30	1-1000
Spatial resolution (m^2) - max	1961	100	50-10000
Model timestep (minutes) - min	12	5	0.02-60
Model timestep (minutes) - max	70	60	2-240
Output speed minutes) - min	42	5	1-300
Output speed (minutes) - max	2577	20	5-21600
Number of ensembles - min	3	1	0-10
Number of ensembles - max	2222352	16	3-2.00E+07

TABLE 15 - BENCHMARKS FOR PERFORMACE CRITERIA - FIRE CHARACTERISATION ACCURACY, TIMESTEP, SPATIAL RESOLUTION, OUTPUT SPEED AND ENSEMBLES SHOWING MEAN, MEDIAN AND RANGE FOR OTHER USER RESPONSES

VARIABLE	MEAN	MEDIAN	RANGE
Rate of spread accuracy (%) - min	53	50	40-70
Rate of spread accuracy (%) - max	92	90	90-95
Area burnt accuracy (%) - min	53	50	40-70
Area burnt accuracy (%) - max	92	90	90-95
Fire behaviour accuracy (%) - min	45	40	30-65
Fire behaviour accuracy (%) - max	83	85	75-90
Impacts accuracy (%) - min	55	50	40-75
Impacts accuracy (%) - max	93	95	90-95
Spatial resolution (m^2) - min	24	10	1-60
Spatial resolution (m^2) - max	393	150	30-1000
Model timestep (minutes) - min	2	1	1-5
Model timestep (minutes) - max	33	30	10-60
Output speed minutes) - min	2	0.2	0.02-5
Output speed (minutes) - max	10	10	1-20
Number of ensembles - min	2	1	1-5
Number of ensembles - max	3357	50	20-10000

Appendix 2. Performance benchmark boxplots with untransformed criteria values

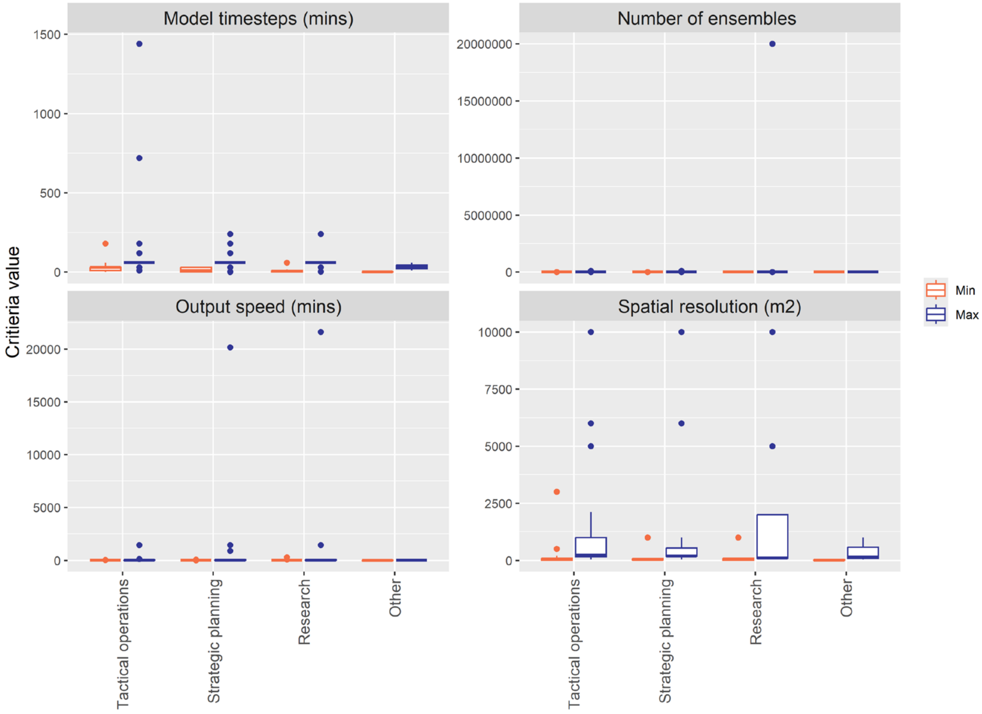

FIGURE 13 – UNTRANSFORMED CRITERIA VALUE WITH OUTLIIERS - PERFORMANCE BENCHMARKS FOR FOUR ASPECTS OF SIMULATOR TOOLS; MODEL TIMESTEPS, NUMBER OF ENSEMBLES, OUTPUT SPEED AND SPATIAL RESOLUTION. ORANGE BOXPLOTS SHOW MINIMUM REQUIREMENTS, BLUE BOXPLOTS SHOW MAXIMUM. RESULTS ARE SEPARATED BY USE CASE. THE LOWER AND UPPER HINGES CORRESPOND TO THE FIRST AND THIRD QUARTILES. WHISKERS EXTEND UP TO 1.5 X THE INTERQUARTILE RANGE.

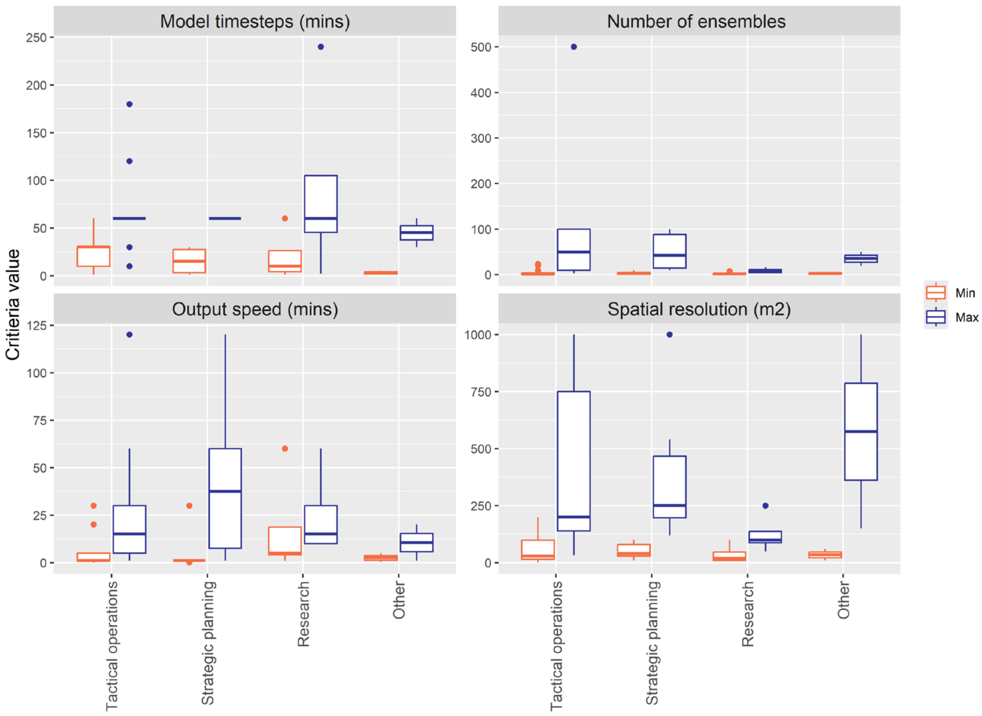

FIGURE 14 – UNTRANSFORMED CRITERIA VALUE WITHOUT OUTLIIERS - PERFORMANCE BENCHMARKS FOR FOUR ASPECTS OF SIMULATOR TOOLS; MODEL TIMESTEPS, NUMBER OF ENSEMBLES, OUTPUT SPEED AND SPATIAL RESOLUTION. ORANGE BOXPLOTS SHOW MINIMUM REQUIREMENTS, BLUE BOXPLOTS SHOW MAXIMUM. RESULTS ARE SEPARATED BY USE CASE. THE LOWER AND UPPER HINGES CORRESPOND TO THE FIRST AND THIRD QUARTILES. WHISKERS EXTEND UP TO 1.5 X THE INTERQUARTILE RANGE.

Appendix 3. Consideration of the nationalisation of fire simulators

Many participants spoke about a national approach to fire simulator development and governance, however there were varied views about if such an approach was ideal or necessary. As an alternative to a completely nationalised approach, many participants preferred a toolkit approach - with some elements nationally coordinated with others locally specific.

Some considerations identified around a nationally co-ordinated approach were:

Positives

- Building of collective resilience.
- Resource sharing, particularly valuable for smaller jurisdictions.

Concerns

- A fear of unequal benefit.
- Decreased jurisdictional or organisational autonomy and control.
- Reduced capacity to respond to localised challenges with simulating.
- Simulators and models not suited to vegetation types (particularly a concern for smaller or remote jurisdictions)
- No appropriate way to standardise data for all jurisdictions.

Other considerations

- Targeted improvement vs broadscale improvement. Do we ensure that the system is acceptable for all users or very good for key users?
- Willingness to invest in a new system can vary between jurisdictions based on the factors outlined in Table 9 that relate to budget, power, risk and capacity to engage in development and implement changes.

Appendix 4. How context dictates fitness-for-purpose

In addition to the benchmarks and priorities for development, creating fit-for-purpose simulators also benefits from case-specific reflection about whether a change is appropriate for users and whether a transition will be able to minimise disruption and maximise positive outcomes. Specific user contexts must then be integrated into simulator design and decision-making and used to facilitate transitions and increase user capacity.

The importance of context in participant discussions was distilled into three key issues outlined in Table 10. These are current circumstances, change implications and change implementation. The lines of questioning associated with each issue collectively consider whether the needs of the user are being met and if not, whether the benefits of a transition will outweigh the necessary effort.

Such lines of questioning may be valuable for systematically consideration of the overarching factors that can influence a whether a simulator is fit-for-purpose.

TABLE 16 - ISSUES TO CONSIDER DURING FIRE SIMULATOR CHANGE PROCESSES, FROM WORKSHOP 1

ISSUE	QUESTIONS
Current circumstances	What is the current system?
	Does the current system function well?
	Why is the change being proposed?
	Who is driving the change?
	Who is the audience for the change?
Change implications	Does the new alternative address the reason for change?
	Is the new alternative fit-for-purpose?
	What are the advantages, costs and trade-offs?
	What is the scale of change?
	What are the long-term benefits that may justify the effort?
Change implementation	What are users' capacities to implement change?
	What are the barriers to implementing change?