Novel method for a posteriori uncertainty quantification in wildland fire spread simulation

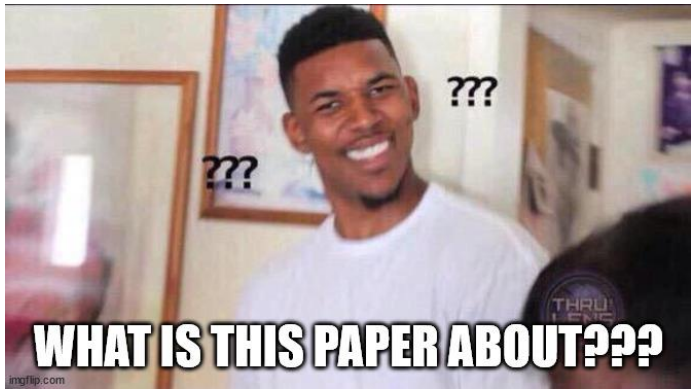
Gautam Dilip Hariharan
Sandeep Bhandari

Supervisor: Daniel Wilmes

Seminar: Uncertainty Quantification in Machine Learning

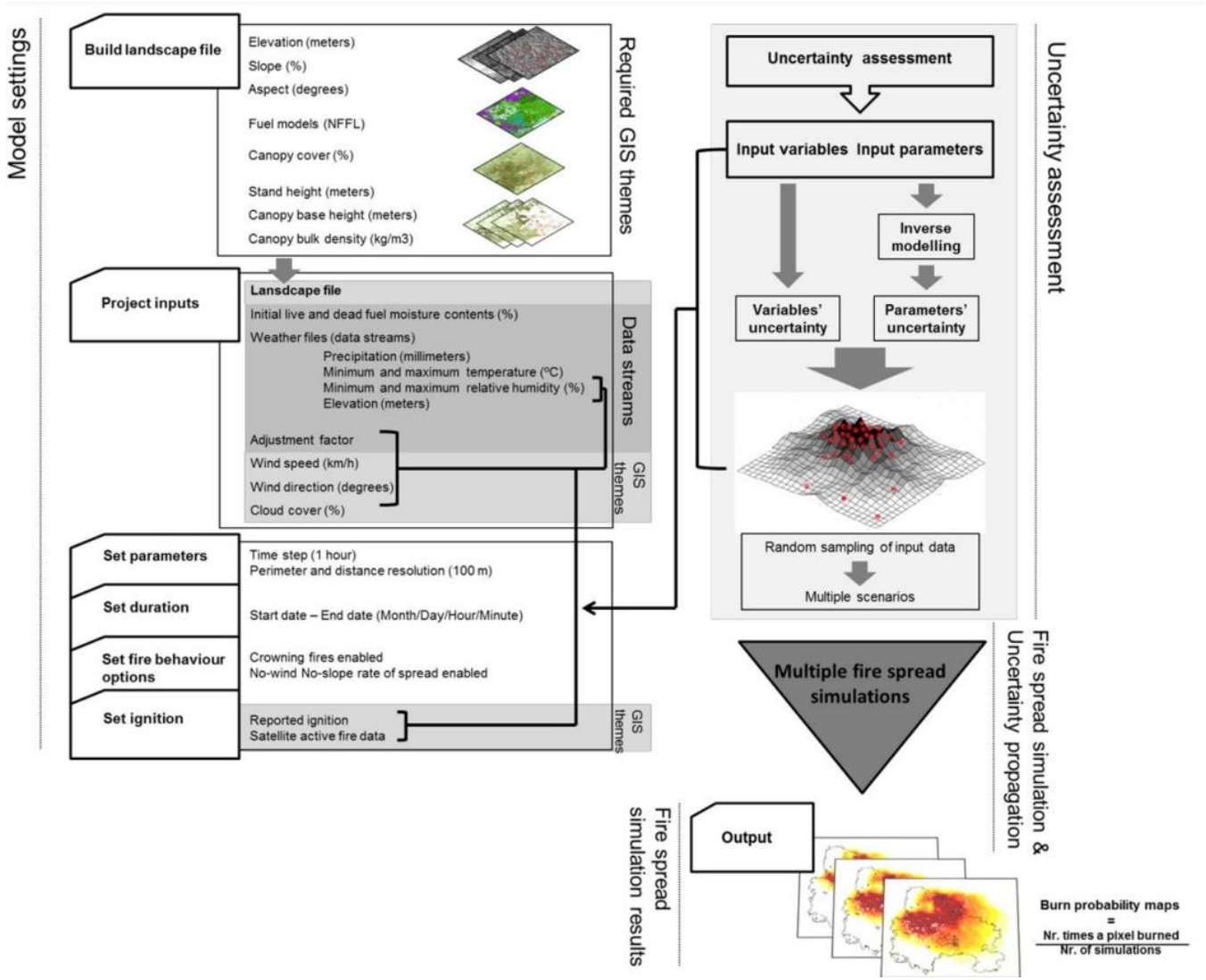
What to expect

- An overview
- Wildland fires and Probabilistic evaluation
- A posteriori uncertainty quantification
- Metropolis Hastings algorithm
- Results
- Conclusion



- Simulations used for prediction of wildland fire spread
- Large uncertainties exist and must be quantified for better accuracy (ensemble forests)
- Generating calibrated ensembles
- Keywords: uncertainty quantification, Metropolis Hastings (MH), Wasserstein distance, Gaussian process

Overview



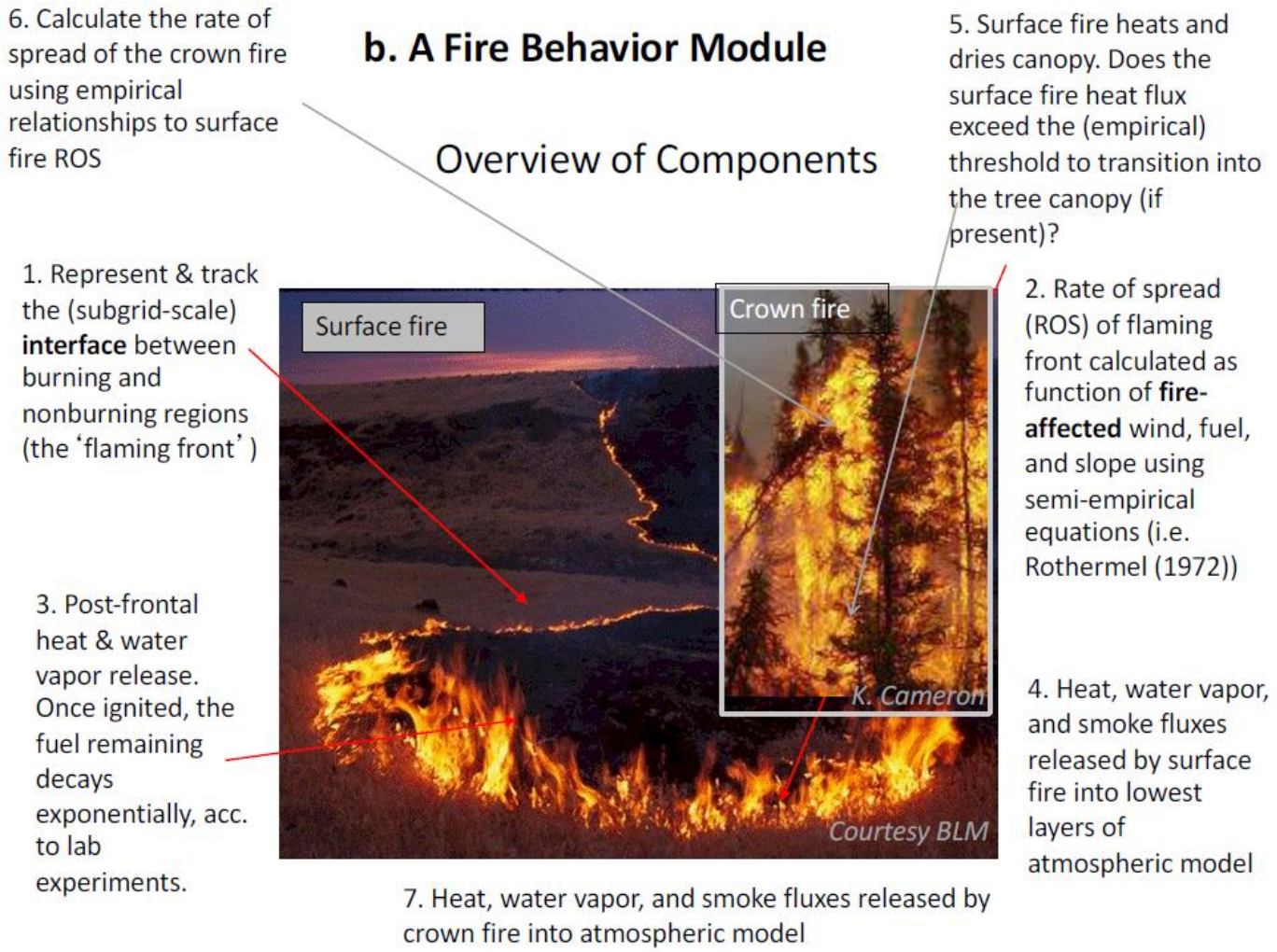
https://springerplus.springeropen.com/track/pdf/10.1186/s40064-016-2842-9.pdf

Input Parameters

- Must have Input File
  - Fuel Grid
    - Is used to define fuel types in the area
    - Must have the same resolution and extent as the DEM file
  - Fuel LookUp Table
    - Defines the fuel types and parameters to be used with the fuel grid.
  - Weather Information
    - Hourly information of temperature,relative humidity,precipitation,wind velocity and direction.
  - Ignition location and time information
    - Supported types are point,line and area ignition
    - Support multiple ignition at different time.
- Optional input files
  - Digital Elevation Model(DEM)
    - Used to calculate the effects on fire spread rate and wind direction
    - Must have the same resolution and extent as the fuel grid file.
  - Physical Features
    - Obstacles that can possibly halt fire propagation(non-fuel areas,water,roads).
  - Wind Grid
    - High Resolution wind velocity and direction grids.
  - Information about the fuel
    - Information about seasonal (green-up phase)
    - Information about the trees(tress height,canopy base height).

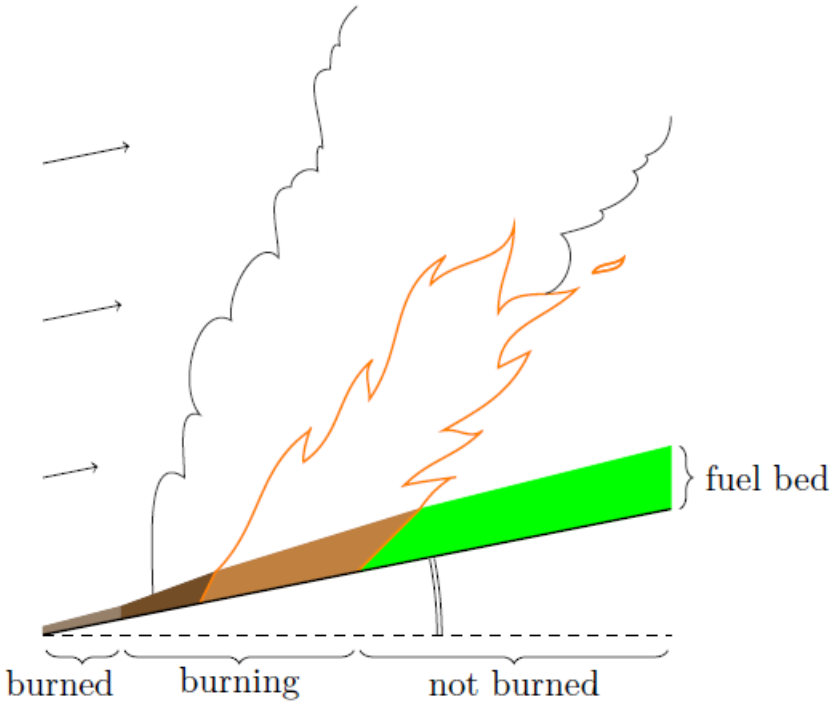
https://springerplus.springeropen.com/track/pdf/10.1186/s40064-016-2842-9.pdf

Fire Behaviour



https://doi.org/10.6028/NIST.SP.1245

Wildlandfire



Packing Ratio

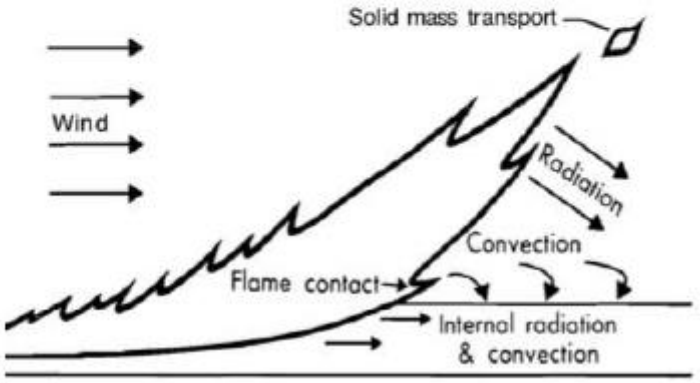
$$\beta_f = \frac{\sigma_f}{h_f \rho_p}.$$

$$R = Ro\,(1 + \phi_w + \phi_s)$$

No-wind spread rate on flat ground

wind factor

slope factor



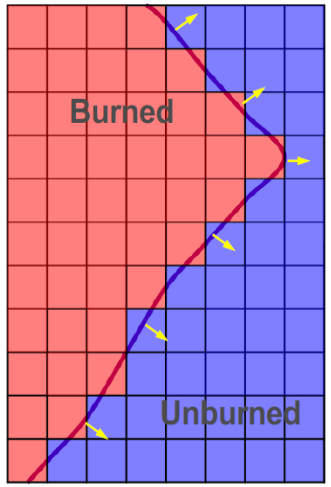
https://www.sciencedirect.com/science/article/abs/pii/S0307904X20304789

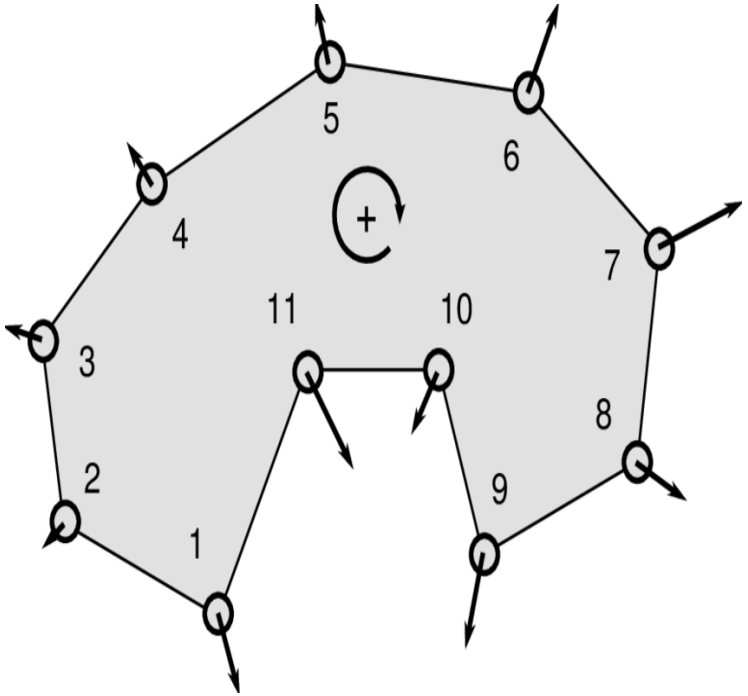
Eulerian Level Set Method

Eulerian Level Set Method

REAX ENGINEERING

- Eulerian: fixed frame of reference such as a grid
- Level set methods are a class of numerical techniques to track surfaces, shapes, or interfaces
- Track curved surfaces on a grid

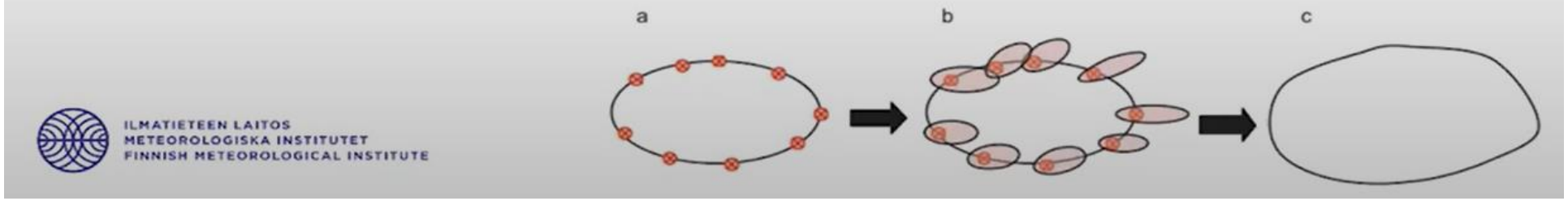




https://doi.org/10.6028/NIST.SP.1245

Fire propagation calculations

- Calculations start by defining a fire perimeter polygon at the ignition location (shape depends on the used ignition type: point, line or area). Figure a.
- After this the propagation calculations are done in a five-stage loop:
  - 1. The propagation of fire is calculated from selected points along the fire perimeter polygon using the FWI/FBP values of the previous propagation loop. Figure b.
  - 2. The newly formed fire perimeter polygon is smoothed. Figure c.
  - 3. If a propagation point has struck with an obstacle (non-fuel cell), the flame length is compared to the width of the obstacle. If it is not able to cross over the obstacle, the propagation point will be turned off. Otherwise, it will cross the boundary and stay active.
  - 4. The perimeter is checked for possible tangles and overlaps with other fire perimeter polygons. Tangles are solved and overlapping perimeters are joined into one.
  - 5. Calculate the FBP/FWI values for the propagation points according to current cell values they are in and restart the loop. If all points have been de-activated, or the simulation time has ended, stop the loop.



ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

https://doi.org/10.6028/NIST.SP.1245

Modeling Uncertainty in Input Data

- Time and Location, Wind Speed highly uncertain
- Solution : Several simulation of fire spread → Perturbation
- Three classes of perturbation : additive, multiplicative and transitive
- additive perturbation, $x_p = x+z$, multiplicative perturbation, $x_p = x \times z$,
- transition perturbation : x →x_p
- Probability distribution to quantify uncertainty →log or log-normal distribution

Input	Unit	Perturbation	σ	Range
Wind direction	°	Additive, global	30	$[-60, 60]$
Wind speed norm	$\mathrm{m\,s^{-1}}$	Multiplicative, global (log)	$0.5 \log 1.5$	$[2/3, 1.5]$
Dead fuel moisture		Multiplicative, global	0.15	$[0.7, 1.3]$

Uncertainty Propagation in Fire Spread Simulation

- Any point xi, b_i (0,1) → prediction

- $q_i = P[B_i = 1]$, B_i follows Bernoulli law of parameters

- Estimate of q_i : $p_i = n_i / n$ (where n_i is the number of simulations for which $b_i = 1$. pi is the MC estimate of q_i and converges to it as n increases)

- p_i is our burn probability and resulting 2D model is our burn probability map.

$$\mathbb{P}\left[|p_i - q_i| \leq \left(\frac{q_i(1-q_i)}{\alpha n}\right)^{1/2}\right] \geq 1 - \alpha,$$

$$\mathbb{P}\left[|p_i - q_i| \leq \left(\frac{\log(2/\alpha)}{2n}\right)^{1/2}\right] \geq 1 - \alpha$$

Probabilistic Evaluation

- Accuracy
  - Measured By Brier Score
  -Ranges between 0 and 1 and negatively oriented
  -BSS(Brier Skill Score) = 1 – BS/BS_{ref}

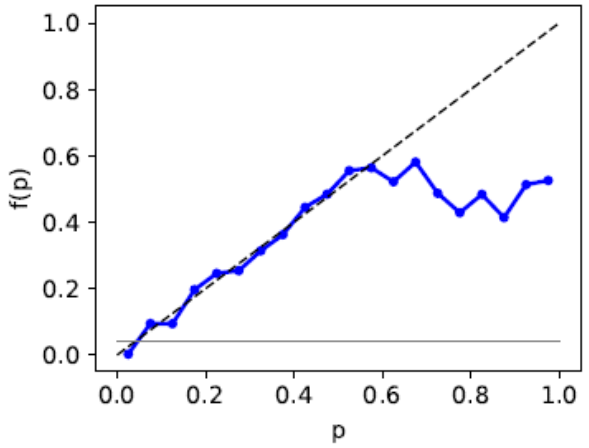
$$BS = \frac{1}{N}\sum_{i=1}^{N}(o_i - p_i)^2.$$

- Reliability and Sharpness
  - A system is reliable if $\forall p \in [0, 1]$, $f(p) = p$.
  - Reliability Diagram (plot of f(p) and p)
  - g(p) as the proportion of events that are assigned a probability p among all evaluated events.
  - The sharpness graph is simply the plot of g(p) against p.

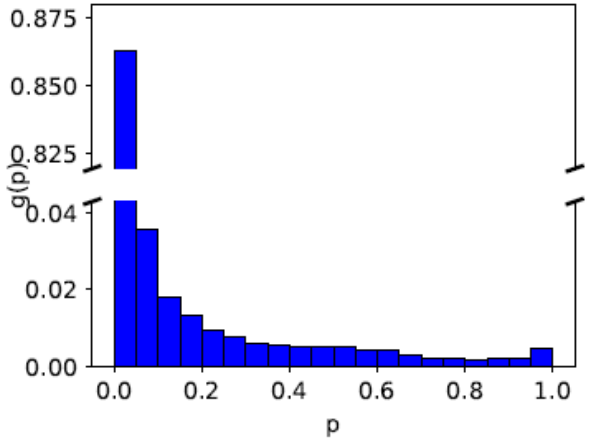
- Probabilistic resolution
  - capacity of the system to yield relative frequencies that are different from the reference probability p_c.
  - Positively correlated
  -Accessed using reliability diagram (deviation between f(p) and p_c.

- Consistency
  -Accessed by Rank Histogram.

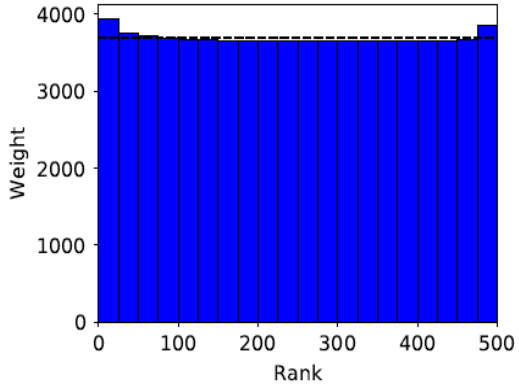
Detailed Case : Calenzana fire

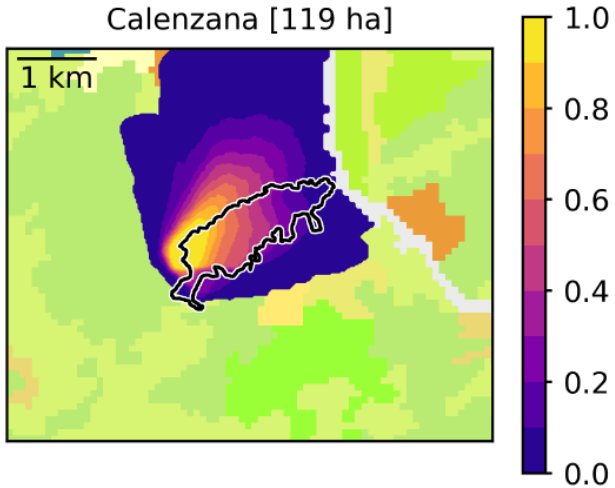


(a) Reliability



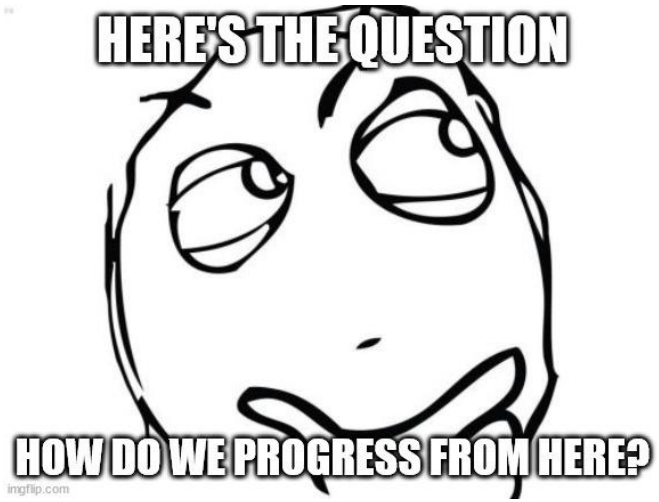
(b) Sharpness





Brier Score = 0.027

https://www.sciencedirect.com/science/article/abs/pii/S0307904X20304789



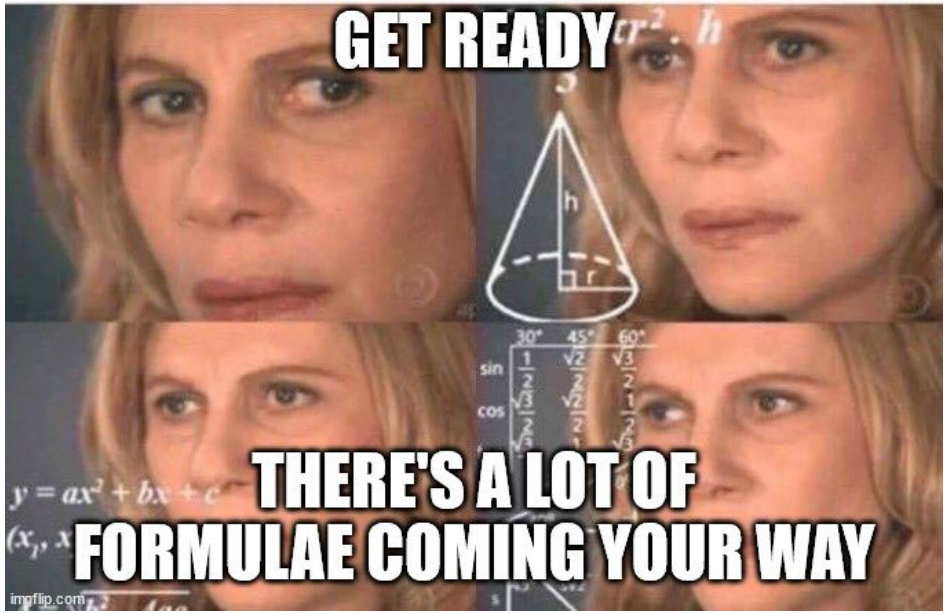
- The goal: To propose and apply a method to calibrate the probability distribution of the inputs of the model based on observed fires
- Solving a problem of inverse uncertainty quantification
- Strategy for calibration of input uncertainty

A posteriori uncertainty quantification

- Sources of uncertainty:
  - Unknown parameters
  - Model inadequacy
  - Observation error

- Bayesian approach proposed for calibration that accounts for different forms of uncertainty.
- Starting point:

$$r(\mathbf{x} - \mathbf{x}') = \exp\left\{ -\sum_{j=1}^{q} \omega_j (x_j - x_j')^2 \right\}.$$


GET READY
THERE'S A LOT OF
FORMULAE COMING YOUR WAY
imgflip.com

Observation of one burned surface

- S_{obs}: observed burned surface
- M: numerical model of fire spread whose inputs may vary according to an input u of d perturbations applied to reference inputs
- S_u: simulated burned surface $(S_u = M(u))$
- g: probability density function
- f: prior density function
- The idea is to obtain g by making the best possible use of f, S_{obs} and M.

Distribution based on Wasserstein distance

- Classical choice for g would be the posterior density function which is obtained according to Bayes's rule:

$$p(\boldsymbol{u}|\mathcal{S}_{\text{obs}}) = \frac{\mathcal{L}(\mathcal{S}_{\text{obs}}|\boldsymbol{u})f(\boldsymbol{u})}{\int \mathcal{L}(\mathcal{S}_{\text{obs}}|\boldsymbol{u})f(\boldsymbol{u})d\boldsymbol{u}}$$

- Where $L(S_{obs}|u)$ would be the likelihood of the observation S_{obs} knowing the perturbation vector u.

- While defining a likelihood for a vector is feasible, this may not be the case for a random surface. Therefore, a calibrated distribution inspired by Bayes' rule may be used where g written in the form:

$$g_{E,\beta}(\boldsymbol{u}) = \frac{e^{-\beta E(\boldsymbol{u})}f(\boldsymbol{u})}{\int e^{-\beta E(\boldsymbol{u})}f(\boldsymbol{u})d\boldsymbol{u}}$$

- Where $\beta>0$ and E is a positive energy function equal to 0 when $S_u = S_{obs}$ and increases with dissimilarity between S_u and S_{obs}.

Distribution based on Wasserstein distance (contd...)

- The present study introduces a novel score that makes use of the Wasserstein distance, which is a metric between probability distributions.
- The square of Wasserstein distance is defined as follows:

$$\mathcal{W}_2^2(\mu,\nu) = \inf\left\{ \int_{\mathbb{R}^q \times \mathbb{R}^q} ||x-y||_2^2 \, d\gamma(x,y) \;\middle|\; \gamma \in \Gamma(\mu,\nu) \right\}$$

- Where $W_2(\mu,\nu)$ is the Wasserstein distance between two probability measures μ *and* ν both defined on R^q, $||.||_2$ is the Euclidian distance and $\Gamma(\mu,\nu)$ is the ensemble of the measures defined on R^q x R^q such that their conditional measure relatively to the first variable is μ and their conditional measure on the second variable is ν.

Distribution based on Wasserstein distance (contd...)

- For comparison between surfaces, it is natural to consider $q = 2$ and choose uniform measures whose support is respectively S_{obs} and S_u for the probability measures μ and ν.
- By making these choices, *E(u)* is defined as:

$$E(\boldsymbol{u}) = \inf_{\gamma} \left\{ \int_{\mathcal{S}_{\text{obs}} \times \mathcal{S}_{\boldsymbol{u}}} ||x - y||_2^2 \, \gamma(x, y) dx dy \;\middle|\; \int_{\mathcal{S}_{\boldsymbol{u}}} \gamma(x, y) dy = \frac{\mathbb{1}(x \in \mathcal{S}_{\text{obs}})}{|\mathcal{S}_{\text{obs}}|}, \; \int_{\mathcal{S}_{\text{obs}}} \gamma(x, y) dx = \frac{\mathbb{1}(y \in \mathcal{S}_{\boldsymbol{u}})}{|\mathcal{S}_{\boldsymbol{u}}|} \right\}$$

- Where **1** is the indicator function, $||.||_2$ is the Euclidian distance and $|S|$ is the surface area of S.

Distribution based on Wasserstein distance (contd...)

- Except for some particular cases, there is no simple analytic formula for the Wasserstein distance.
- This led us to consider a discrete approximation of *E(u)* instead, which can be obtained numerically via a discretization of the PDF's by a sum of Dirac delta distributions.
- From this point, *E(u)* is now defined as:

$$E(\boldsymbol{u}) = \inf_{\gamma} \left\{ \int_{\mathcal{S}_{\text{obs}} \times \mathcal{S}_{\boldsymbol{u}}} ||x-y||_2^2 \, \gamma(x,y) dx dy \;\middle|\; \int_{\mathcal{S}_{\boldsymbol{u}}} \gamma(x,y) dy = \frac{1}{J} \sum_{j=1}^{J} \delta_{x_j}(x), \right.$$
$$\left. \int_{\mathcal{S}_{\text{obs}}} \gamma(x,y) dx = \frac{1}{K} \sum_{k=1}^{K} \delta_{y_k}(y) \right\}$$

  - Where δ_x is the Dirac delta distribution at point $x \in R^2$, and each x_j belongs to S_{obs}, whereas y_k belongs to S_u

Distribution based on Wasserstein distance (contd...)

- In this discrete setting, the admissible distributions γ can be represented by a matrix of size J x K where each cell γ_{jk} is positive and indicates the "probability mass" that is transferred from x_j to y_k.
- In this case, the infimum of the equation in the previous slide is reached and is the solution of the following linear programming problem:

$$\min_{\gamma_{j,k}} \sum_{j=1}^{J} \sum_{k=1}^{K} \gamma_{jk} ||x_j - y_k||_2^2$$

$$\text{subject to } \gamma_{jk} \geq 0, \sum_j \gamma_{jk} = \frac{1}{K}, \text{ and } \sum_k \gamma_{jk} = \frac{1}{J}.$$

- Which is also referred to as Earth Mover's Distance. It is known from graph theory that the optimal γ is a sparse matrix that has at most J + K – 1 non-zero cells.

Distribution based on Wasserstein distance (contd...)

- An issue in the denominator of $g_{E,\beta}$ (from the equation in slide 18) which is an intractable high-dimensional integral.
- But this integral does not depend on the perturbation vector u, so for a given β, the PDF is known up to some constant factor.
- Another question arises: how do we draw samples from the distribution when it is known up to a factor?


THE METROPOLIS
HASTINGS ALGORITHM
imgflip.com

Emulation

- Lot of iterations of MH algorithm required to obtain a sufficiently large sample, takes too much time
- Emulator ($\tilde{E}$) used to speed up MH
- $\tilde{E}(u)$ provides good approximation of $E(u)$ while being considerably fast to compute.

Gaussian Process Modelling

- Emulation method, also called kriging
- In statistics, Gaussian process emulator is one name for a general type of statistical model that has been used in contexts where the problem is to make maximum use of the outputs of a complicated computer-based simulation model.

Design of experiments and error metrics

- Inputs of training sample are obtained via Latin hypersquare sample with optimized discrepancy
- Complementary test sample generated to evaluate approximation error of emulator far from training points (obtained with an algorithm for an optimal validation design)
- Based on test sample, several error metrics used to evaluate emulator, here Mean Absolute Error (MAE), standardized mean square error (SMSE) and Q_2 metric
- As the error of the emulator gets lower, MAE gets closer to 0 and Q_2 metric gets closer to 1

Extension to several fire cases

- Considering K fire cases, it is possible to compute the energy functions that correspond to each fire ($E_1,\ldots,E_K$)
- An issue arises when the variations of one energy function are much higher than for the other fires in which case the variations of pseudo-likelihood will mostly be determined by that one energy function and the calibrated distribution will mostly be representative of that one fire at the expense of the other observations
- How do we fix this? Weigh each fire depending on the values taken by $E_K(u)$ and define the energy functions as the weighted sum of squared Wasserstein distances:

$$E(\boldsymbol{u}) = \sum_{k=1}^{K} w_k E_k(\boldsymbol{u}),$$

where the weights are defined using all points from the training dataset:

$$w_k = \frac{n}{\sum_{i=1}^{n} E_k(\boldsymbol{u}^i)}.$$

Sampling from a calibrated distribution

- As already explained before, emulator used instead of energy function to run the algorithm in reasonable time
- m: number of chains
- n: number of samples per chain
- $u_{i,j}$: i^{th} element of the j^{th} chain
- $u_{c,j}$: candidate
- This algorithm is motivated by the convergence diagnosis
- Recommended to choose $u_{1,1},\ldots,u_{1,m}$ quite far from each other
- Based on chains returned by the MH algorithm, the matrices B/n (size d) and W are computed as:

$$B/n = \frac{1}{m-1}\sum_{j=1}^{m}(\bar{\boldsymbol{u}}_j - \bar{\boldsymbol{u}})(\bar{\boldsymbol{u}}_j - \bar{\boldsymbol{u}})^T,$$

$$W = \frac{1}{m(n-1)}\sum_{j=1}^{m}\sum_{i=1}^{n}(\boldsymbol{u}_{i,j} - \bar{\boldsymbol{u}}_j)(\boldsymbol{u}_{i,j} - \bar{\boldsymbol{u}}_j)^T,$$

- $\bar{u}_j$ is the sample mean of the j^{th} chain and $\bar{u}$ is the sample mean over all chains

Sampling from a calibrated distribution (contd...)

- The metric used for analysing convergence is:

$$\hat{R}^d = \frac{n-1}{n} + \left(\frac{m+1}{m}\right)\lambda_1,$$

- Where λ_1 is the largest eigenvalue of the symmetric, positive definite matrix $W^{-1}B/n$
- At convergence, $\hat{R}^d$ tends to 1 and one may consider that a sufficient number of MH iterations has been carried out if $\hat{R}^d < 1.1$ for the second half of the chains
- From this, it follows that the set comprising the second half of all m chains constitutes a representative sample of the target distribution when $\hat{R}^d < 1.1$

Metropolis Hastings algorithm (Sampling from the calibrated distribution)

Define m, n, and an instrumental distribution of PDF $q : \boldsymbol{u} \mapsto q(\boldsymbol{u}|\boldsymbol{v})$
for $j = 1, ..., m$ **do**
  Choose a starting point $\boldsymbol{u}_{1,j}$
  for $i = 2, ..., n$ **do**
    Sample a candidate $\boldsymbol{u}_{c,j} \sim q(.|\boldsymbol{u}_{i-1,j})$
    Compute the ratio

$$\tau = \frac{g_{\widetilde{E},\beta}(\boldsymbol{u}_{c,j})\ q(\boldsymbol{u}_{i-1,j}|\boldsymbol{u}_{c,j})}{g_{\widetilde{E},\beta}(\boldsymbol{u}_{i-1,j})\ q(\boldsymbol{u}_{c,j}|\boldsymbol{u}_{i-1,j})} = \frac{e^{-\beta\widetilde{E}(\boldsymbol{u}_{c,j})}\ f(\boldsymbol{u}_{c,j})\ q(\boldsymbol{u}_{i-1,j}|\boldsymbol{u}_{c,j})}{e^{-\beta\widetilde{E}(\boldsymbol{u}_{i-1,j})}\ f(\boldsymbol{u}_{i-1,j})\ q(\boldsymbol{u}_{c,j}|\boldsymbol{u}_{i-1,j})}$$

    if $\tau \geq 1$ **then**
      [Accept the candidate]
      $\boldsymbol{u}_{i,j} \leftarrow \boldsymbol{u}_{c,j}$
    else
      [Accept the candidate with probability τ]
      Sample $p \sim \mathcal{U}(0, 1)$
      if $p \leq \tau$ **then**
        $\boldsymbol{u}_{i,j} \leftarrow \boldsymbol{u}_{c,j}$
      else
        $\boldsymbol{u}_{i,j} \leftarrow \boldsymbol{u}_{i-1,j}$
      end if
    end if
  end for
end for
return $(\boldsymbol{u}_{1,j}, ..., \boldsymbol{u}_{n,j})_{j=1,...,m}$

Fire spread simulation

- Open source fire spread solver ForeFire is used
- Input variables of ROS model subject to perturbation:
  - Fuel moisture content of dead fuel
  - Surface volume ratio
  - Heat content
  - The fuel load
  - The particle density
  - The fuel bed depth
  - “effective” wind speed in the direction of fire spread
- Some assumptions:
  - Mineral damping coefficient = 1
  - Fuel mineral content is negligible (net initial fuel loading is equal to fuel load)

Fire spread simulation (contd...)

- Additionally, to account for wind speed at mid-height of the flame being lower than that of the prediction, a 0.4 factor in ROS computations is applied to *W* so that W_S = *0.4 W.n*
- The scheme used to advance the markers of the fire front is based on a first-order approximation
- Considering a marker that is located at x_i at time t_i, with its normal to the front denoted as $\boldsymbol{n}_i$ (oriented toward the unburned area), its next location is determined by:

$$x_{i+1} = x_i + \delta l \; \boldsymbol{n}_i$$

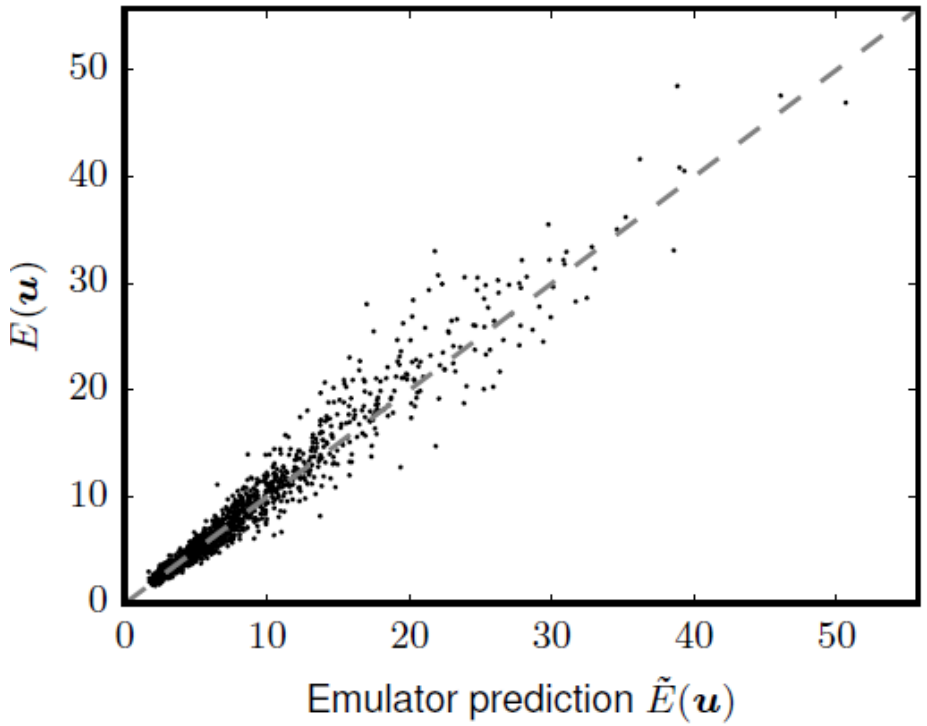
- The advance in time depends on ROS_i, the ROS computed with the values of the environmental inputs at location x_i and time t_i, as follows:

$$t_{i+1} = t_i + \frac{\delta l}{ROS_i}$$

Application to seven Corsican wildland fires

- *K = 7 fires*
- Previous Study Ensembles to as **"reference** ensembles"[1]
- Emulator training size of **4000**,test sample size of **2000**
- Resolution = 20m for small surface and 40m for large surface
- Package ***ot*** *from the Python toolbox* POT[2]
- *MH algorithm is applied for different distribution with different values of* **β={1/20,1/7,1/4,1/2,1}**
- For each value of **β** , *n = 150000 iterations* ***are*** **carried out for** ***m = 8 chains***
- The distribution is then truncated to the perturbation range.
- Take the latter half of chain of MH→ *m × n/2 = 600000 for each* **β**
- Ensembles of wildland fire simulation is carried out → **calibrated ensembles**
- size of a calibrated ensemble ranges between 2000 and 10000
- ensemble generation based on prior distribution→ **prior ensembles**

Results

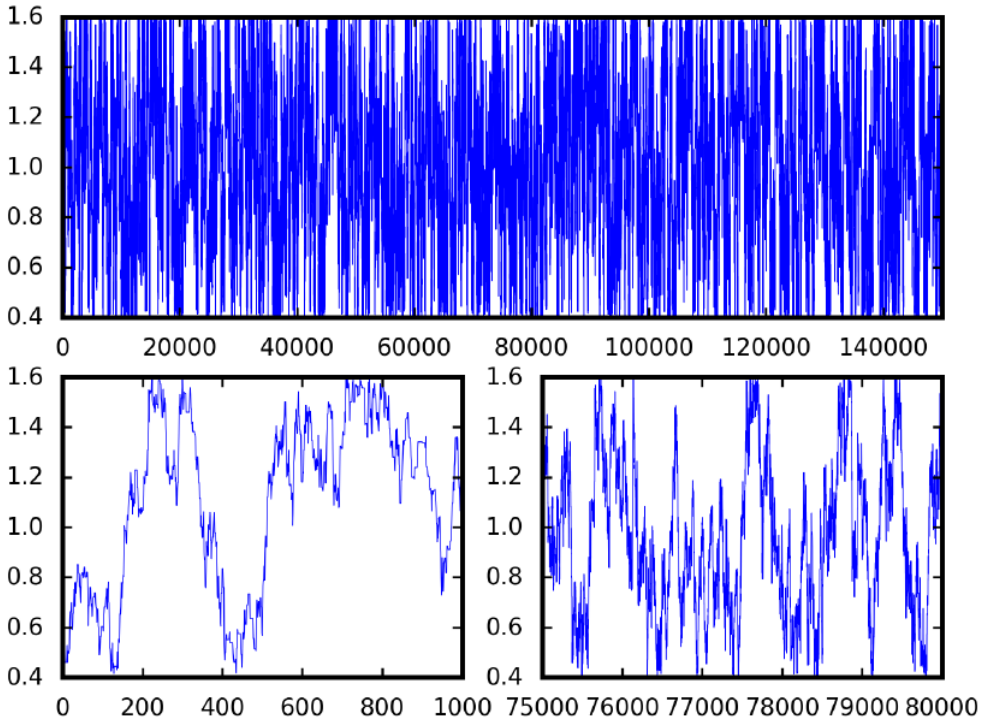


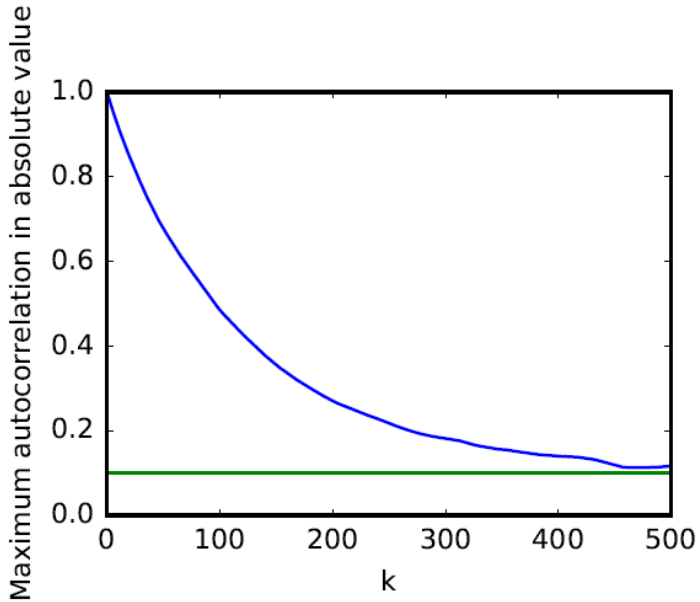
MAE = 0.73
Q_2 = 95.3%

- Use of Logarithm favoured
- Without logarithm: MAE = 0.97, Q_2 = 93.2%
- 0.6s for one energy function
- 150000 iteration of MH algorithm-> more than a day

https://www.sciencedirect.com/science/article/abs/pii/S0307904X20304789

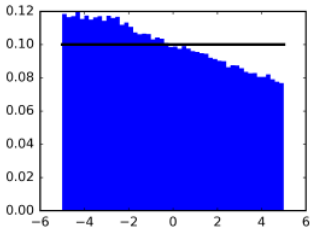
Results : Calibrated Distribution



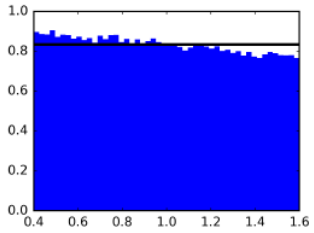


https://www.sciencedirect.com/science/article/abs/pii/S0307904X20304789

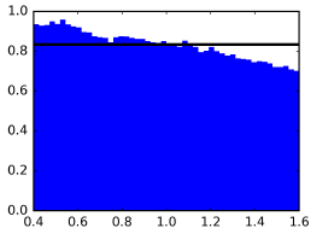
Results : Calibrated Distribution



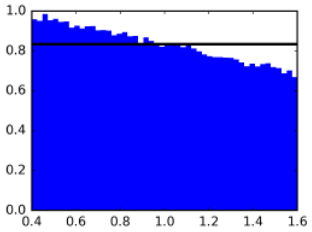
(a) ΔH



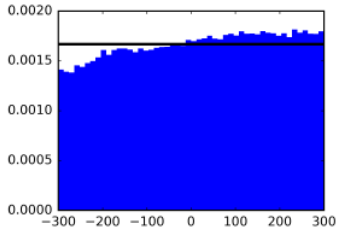
(b) σ_f of fuel type 311



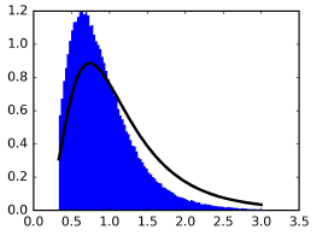
(c) S_v of fuel type 311



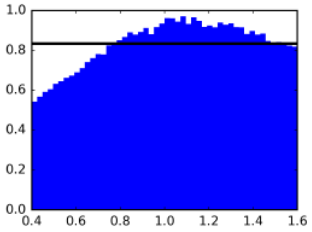
(d) h of fuel type 311



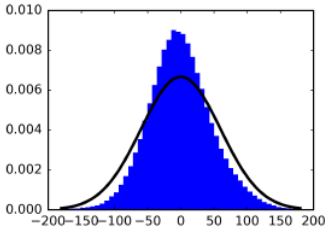
(e) ρ_p



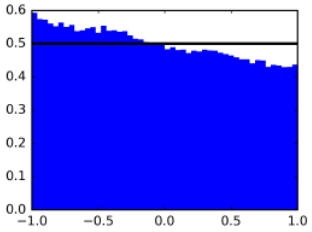
(f) Wind speed norm



(g) m_c



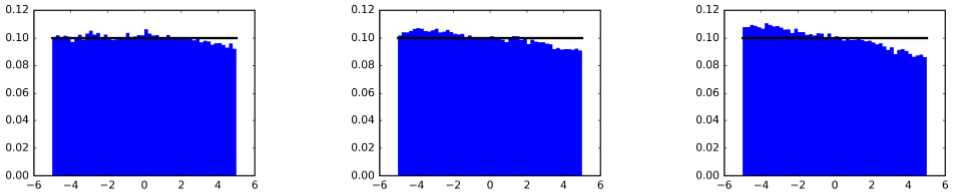
(h) Wind direction



(i) Time of fire end

https://www.sciencedirect.com/science/article/abs/pii/S0307904X20304789

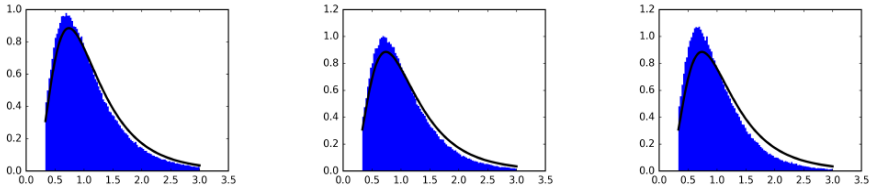
Results : Calibrated Distribution



(a) $\beta = 1/10$ (b) $\beta = 1/7$ (c) $\beta = 1/4$

(d) $\beta = 1/2$ (e) $\beta = 1$ (f) $\beta = 2$

Marginal calibrated distribution of ΔH *for different values of* **β**



(a) $\beta = 1/10$ (b) $\beta = 1/7$ (c) $\beta = 1/4$

(d) $\beta = 1/2$ (e) $\beta = 1$ (f) $\beta = 2$

Marginal calibrated distribution of wind speed norm *for different values of* **β**

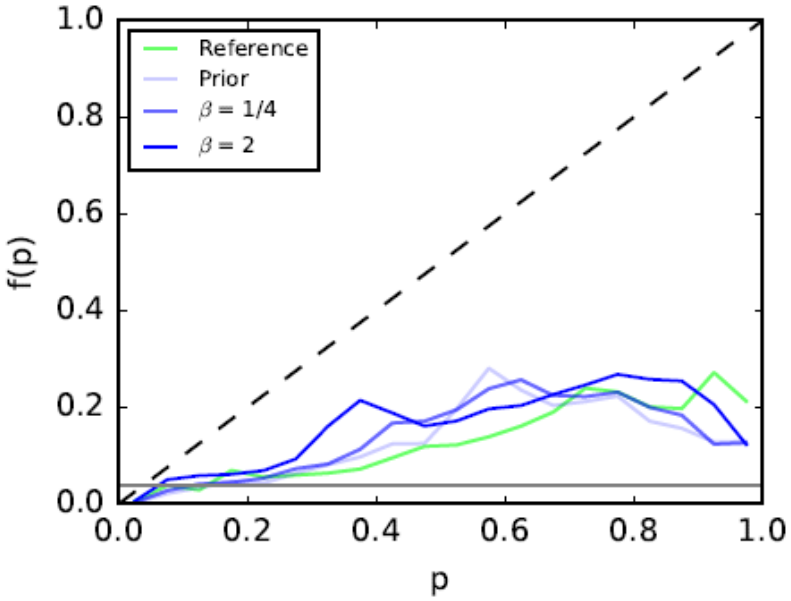
Results : Ensemble Evaluation

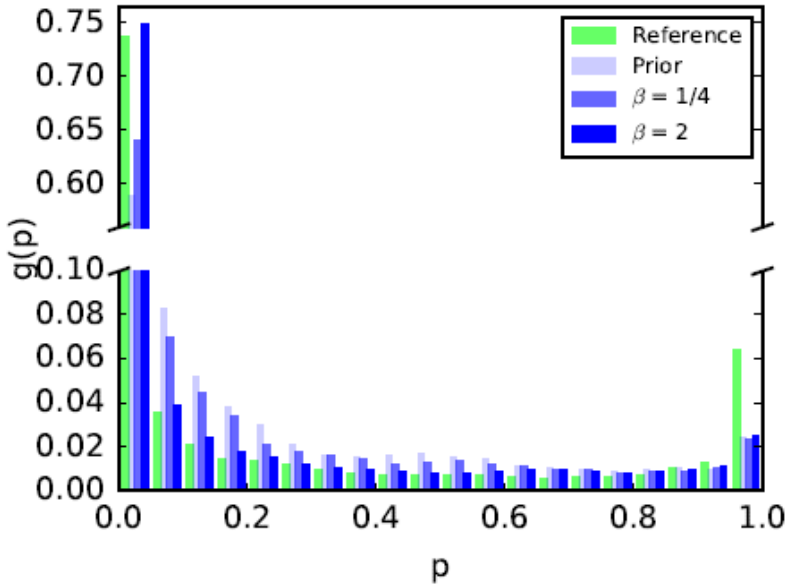
Fire name (ensemble size)	Reference	Prior	$\beta = 1/10$	$\beta = 1/7$	$\beta = 1/4$	$\beta = 1/2$	$\beta = 1$	$\beta = 2$
Calenzana (10000)	0.269	0.291	0.304	0.308	0.309	**0.314**	0.308	0.284
Chiatra (10000)	0.324	**0.386**	0.385	0.379	0.371	0.358	0.342	0.325
Ville di Paraso (2000)	0.021	0.168	0.179	0.182	**0.189**	0.188	0.176	0.168
Sant'Andrea di Cotone (5000)	0.190	0.408	0.429	0.442	0.454	0.468	0.485	**0.494**
Olmeta di Tuda (2000)	0.063	0.187	0.230	0.219	0.278	0.322	0.378	**0.451**
Nonza (4000)	-5.323	-3.089	-3.124	-3.133	-3.124	**-3.044**	-3.057	-3.053
Ghisoni (2000)	-9.986	-10.273	-9.831	-9.851	-9.333	-9.018	-8.638	**-8.332**
Global	-1.609	-1.332	-1.266	-1.269	-1.191	-1.135	-1.080	**-1.033**

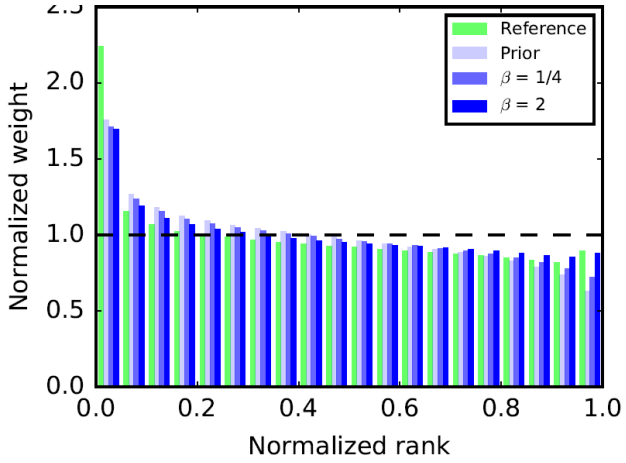
Fire name (ensemble size)	Reference	Prior	$\beta = 1/10$	$\beta = 1/7$	$\beta = 1/4$	$\beta = 1/2$	$\beta = 1$	$\beta = 2$
Calenzana (10000)	8	6	5	3	2	1	3	7
Chiatra (10000)	8	1	2	3	4	5	6	7
Ville di Paraso (2000)	8	6	4	3	1	2	5	6
Sant'Andrea di Cotone (5000)	8	7	6	5	4	3	2	1
Olmeta di Tuda (2000)	8	7	5	6	4	3	2	1
Nonza (4000)	8	4	5	7	5	1	3	2
Ghisoni (2000)	7	8	5	6	4	3	2	1
Sum	55	39	32	33	24	18	23	25
Overall ranking	8	7	5	6	3	1	2	4

https://www.sciencedirect.com/science/article/abs/pii/S0307904X20304789

Results : Ensemble Evaluation

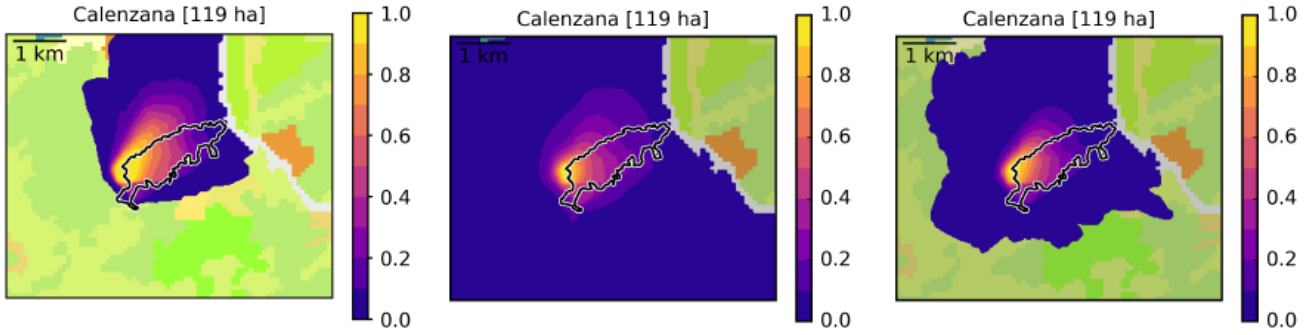




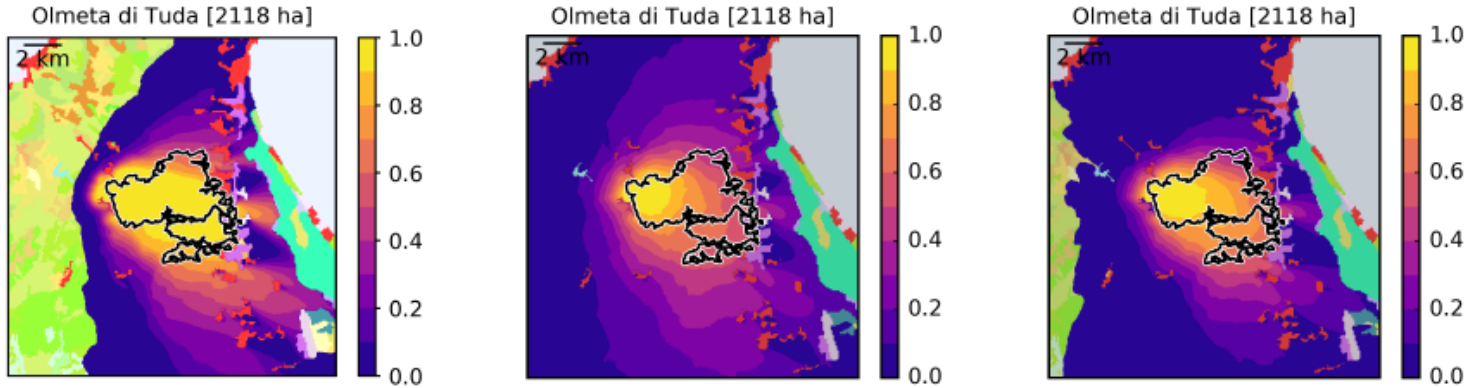


https://www.sciencedirect.com/science/article/abs/pii/S0307904X20304789

Results : Ensemble Evaluation



(a) Reference (b) Prior (c) Calibrated, $\beta = 2$



(a) Reference (b) Prior (c) Calibrated, $\beta = 2$

https://www.sciencedirect.com/science/article/abs/pii/S0307904X20304789

Conclusion

- Here's what we know so far:
  - Led to generation of calibrated ensembles (input distributions defined by posterior PDF) with a pseudo likelihood function that involves the Wasserstein distance between simulated and burned surfaces
  - Gaussian process emulator was built to obtain calibrated sample because of high dimensionality and computational requirements
  - Emulation showed good accuracy ($Q_2 > 95\%$)
  - Calibration was successful in modifying the probability distribution of the input so that the fire spread predictions have better overall accuracy
  - Safe to assume that increasing β lead to distributions that favour lower ROS
  - Best overall BSS ranking for $\beta = ½$, not best globally but very good one for most fires

For consideration

- Could take into account other sources of uncertainty in calibration like model error—not really straightforward because of the nature of model input
- All large fires for one season and from one region were chosen. No guarantee that there will still be an overall improvement if other fires are included
- More fires in the training sample would provide more information, should limit overfitting
- Improving prediction accuracy is crucial because many important parts of our ecosystem are endangered here

Further research

Main research perspective is now to combine these calibrated ensembles with models for probability of ignition and values at stake to assess next day wildfire risk, which is relevant to fire managers, and help in the decision of firefighting actions and fire prevention planning.

References

[1] F. Allaire, J.-B. Filippi, V. Mallet, Generation and evaluation of an ensemble of wildland fire simulations, International Journal of Wildand Fire 29 (2) (2020) 160–173. doi:10.1071/WF19073.

[2] R. Flamary, N. Courty, POT Python Optimal Transport library (2017). URL https://pythonot.github.io/


THANK YOU
QUESTIONS?
imgflip.com