A CFD-based multi-fidelity surrogate model for prediction of flow parameters in a ventilated room

N. Morozova[1], F.X. Trias[1], V. Vanovskiy[2], C. Oliet[1]
and E. Burnaev[2]

1 - Heat and Mass Transfer Technological Centre (CTTC),
Universitat Politècnica de Catalunya (UPC)
2 - I was asked to remove this affiliation

The 8th European Congress on Computational Methods in
Applied Sciences and Engineering
ECCOMAS Congress 2022
5-–9 June 2022, Oslo, Norway

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Outline

Introduction

Simulations of indoor environment. State of the art

- Modern heating ventilation and air conditioning (HVAC) systems are required to maintain a trade-off between maximizing human thermal comfort and minimizing energy consumption in buildings.
- The air distribution in buildings is usually evaluated either by simplified reduced-order models or by CFD.
- Simplified models provide very rapid predictions but offer limited information due to assumptions required.
- CFD simulations are computationally too expensive.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Introduction

Simulations of indoor environment. State of the art

- Modern heating ventilation and air conditioning (HVAC) systems are required to maintain a trade-off between maximizing human thermal comfort and minimizing energy consumption in buildings.
- The air distribution in buildings is usually evaluated either by simplified reduced-order models or by CFD.
- Simplified models provide very rapid predictions but offer limited information due to assumptions required.
- CFD simulations are computationally too expensive.

Introduction

Simulations of indoor environment. State of the art

- Modern heating ventilation and air conditioning (HVAC) systems are required to maintain a trade-off between maximizing human thermal comfort and minimizing energy consumption in buildings.
- The air distribution in buildings is usually evaluated either by simplified reduced-order models or by CFD.
- Simplified models provide very rapid predictions but offer limited information due to assumptions required.
- CFD simulations are computationally too expensive.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Introduction

Simulations of indoor environment. State of the art

- Modern heating ventilation and air conditioning (HVAC) systems are required to maintain a trade-off between maximizing human thermal comfort and minimizing energy consumption in buildings.
- The air distribution in buildings is usually evaluated either by simplified reduced-order models or by CFD.
- Simplified models provide very rapid predictions but offer limited information due to assumptions required.
- CFD simulations are computationally too expensive.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Introduction

Objectives of the study

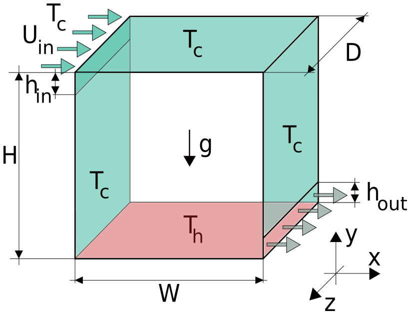
1. Explore the capabilities of the surrogate modeling as a cheaper alternative to CFD for indoor environmental applications.
2. Reduce the computational cost of the dataset generation by implementing multi-fidelity approach.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Physical problem

Three-dimensional mixed convection in a ventilated cavity

- 3D ventilated room filled with air, where the floor is heated, side walls are cold, front and rear walls are periodic. The inlet flow is cold.
- Experimental parameters:
  1. Cavity width aspect ratio A_W;
  2. Froude number $Fr = U_{in}/U_{buo}$;
  3. Rayleigh number $Ra_H = g\beta\Delta TH^3/(\nu\alpha)$.



Problem description

Configuration of the experiments

Experimental parameters

A_w	$0.25, 0.50, 1.00, 2.00, 4.00$
Ra_H	$1.5 \times 10^8, 6.0 \times 10^8, 2.4 \times 10^9, 9.6 \times 10^9$
Fr	$0.15, 0.20, 0.25, ..., 0.55, 0.60, 0.70, ..., 1.50, 1.60$

CFD simulations

Total number of low-fidelity (LF) simulations - 360
Total number of high-fidelity (HF) simulations - 240
Total number of simulations - 600

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Problem description

Input-output parameters

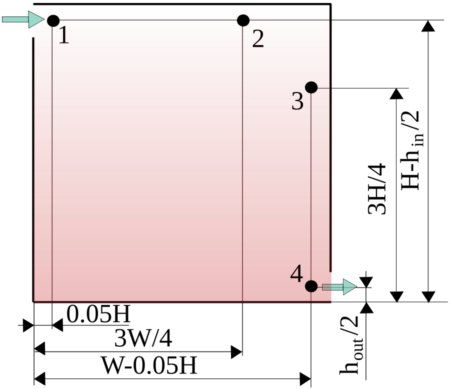
Input parameters

1. A_w
2. $T_1, ..., T_4$
3. $V_1, ..., V_4$

Output parameters

1. Nusselt number on the hot wall - $< Nu >$
2. Average enstrophy - $< \Omega >$

1



[1]N. Morozova, F. X. Trias, R. Capdevila, E. Schillaci and A. Oliva. *A CFD-based surrogate model for predicting flow parameters in a ventilated room using sensor readings*. Energy Build, **266**:112146, 2022.

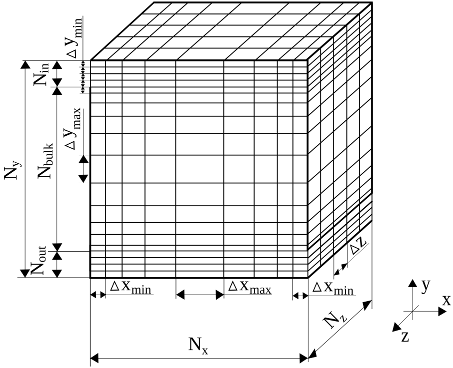
Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Numerical methods

CFD simulations

Automatic mesh generation

1. $\Delta x_{max} = \Delta y_{max} \approx C_r \eta_{Grö}$
2. $\Delta y_{min} \approx C_r \frac{h_{in}}{2} Re_\tau^{-1}$
3. $\Delta x_{min} = 4\Delta y_{min}$
4. $\gamma_x, N_x, \gamma_y, N_b ulk$ - are found iteratively using the hyperbolic tangent function
5. $N_z = 1.1D/\Delta y_{max}$



Grid resolutions

- $C_r = 3$ - high-fidelity simulations
- $C_r = 6$ - low-fidelity simulations

Numerical methods

Surrogate models - 1

- The work is based on the Gaussian process regression (GPR).
- All of the models use an open-source library scikit-learn.

Single-fidelity models

- **HF-GPR** - a model trained only high-fidelity (HF) data;
- **LF-GPR** - a model trained only on low-fidelity (LF) data.
- **HFLF-GPR** - a model trained on a mix of LF and HF data without distinguishing the data fidelity.

Numerical methods

Surrogate models - 1

- The work is based on the Gaussian process regression (GPR).
- All of the models use an open-source library scikit-learn.

Single-fidelity models

- **HF-GPR** - a model trained only high-fidelity (HF) data;
- **LF-GPR** - a model trained only on low-fidelity (LF) data.
- **HFLF-GPR** - a model trained on a mix of LF and HF data without distinguishing the data fidelity.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Numerical methods

Surrogate models - 1

- The work is based on the Gaussian process regression (GPR).
- All of the models use an open-source library scikit-learn.

Single-fidelity models

- **HF-GPR** - a model trained only high-fidelity (HF) data;
- **LF-GPR** - a model trained only on low-fidelity (LF) data.
- **HFLF-GPR** - a model trained on a mix of LF and HF data without distinguishing the data fidelity.

Numerical methods

Surrogate models - 2

GPR with linear correction (LCGPR)

1. Train a single-fidelity model on LF data and test it on HF data - $M_l(\mathbf{X}_l, \mathbf{y}_l) = \tilde{\mathbf{y}}_h$.
2. Estimate an error $(\Delta \mathbf{y}_h)$ between the test results and the actual high-fidelity data and train a linear regression model $\Delta M(\mathbf{X}_h, \Delta \mathbf{y}_h)$ to predict this error.
3. Correct the predictions of low-fidelity surrogate model for the the step one - $M_l(\mathbf{X}_h)$ using the error correction model from the step two.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Numerical methods

Surrogate models - 2

GPR with linear correction (LCGPR)

1. Train a single-fidelity model on LF data and test it on HF data - $M_l(\mathbf{X}_l, \mathbf{y}_l) = \tilde{\mathbf{y}}_h$.
2. Estimate an error $(\Delta\mathbf{y}_h)$ between the test results and the actual high-fidelity data and train a linear regression model $\Delta M(\mathbf{X}_h, \Delta\mathbf{y}_h)$ to predict this error.
3. Correct the predictions of low-fidelity surrogate model for the the step one - $M_l(\mathbf{X}_h)$ using the error correction model from the step two.

Numerical methods

Surrogate models - 2

GPR with linear correction (LCGPR)

1. Train a single-fidelity model on LF data and test it on HF data - $M_l(\mathbf{X}_l, \mathbf{y}_l) = \tilde{\mathbf{y}}_h$.
2. Estimate an error $(\Delta\mathbf{y}_h)$ between the test results and the actual high-fidelity data and train a linear regression model $\Delta M(\mathbf{X}_h, \Delta\mathbf{y}_h)$ to predict this error.
3. Correct the predictions of low-fidelity surrogate model for the the step one - $M_l(\mathbf{X}_h)$ using the error correction model from the step two.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Numerical methods

Surrogate models - 3

Multi-fidelity GPR (MFGPR) or co-kriging

It estimates for a poorly sampled variable $\mathbf{y}_l(\mathbf{x})$ with the help of a well-sampled variable $\mathbf{y}_h(\mathbf{x})$:

$$\mathbf{y}_l(\mathbf{x}) = f_l(\mathbf{x}) + \epsilon_l$$

$$\mathbf{y}_h(\mathbf{x}) = \rho y_l(\mathbf{x}) + y_d(\mathbf{x})$$

Numerical methods

Metrics

Mean relative prediction error (MRE)

$$MRE(\phi) = \frac{1}{N}\sum_{i=1}^{N} \frac{|\phi_{CFD} - \phi_{SM}|}{|\phi_{CFD}|}$$

We assume that less than **10% MRE** is acceptable for this model.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

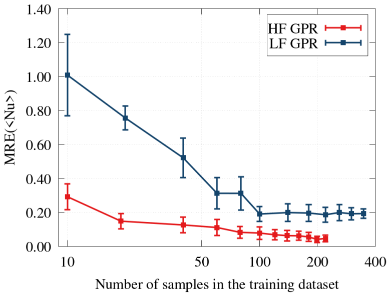
Results

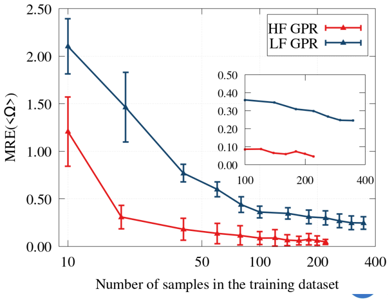
Single-fidelity models

Average computational cost

2700 core-hours - a high-fidelity simulation

285 core-hours - a low-fidelity simulation

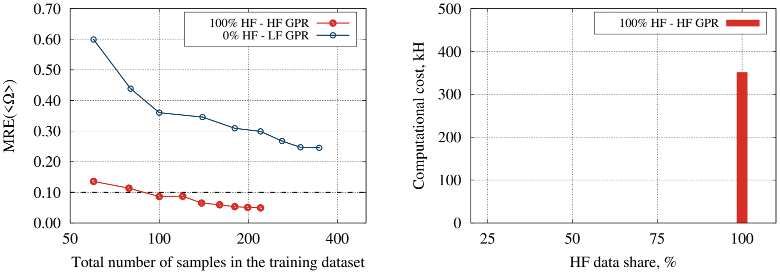




Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

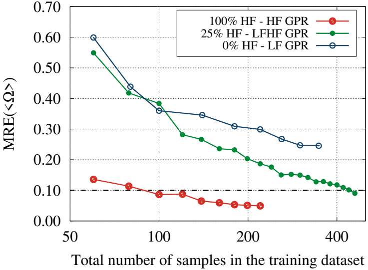
Results

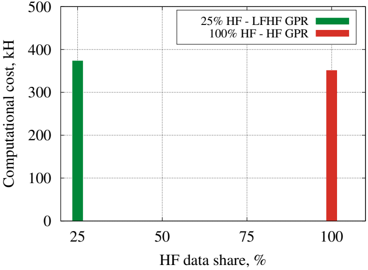
HFLF GPR model



Results

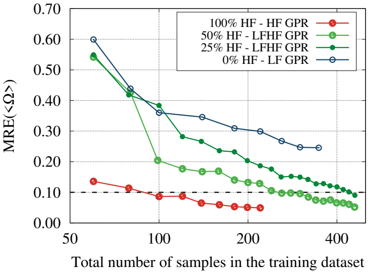
HFLF GPR model

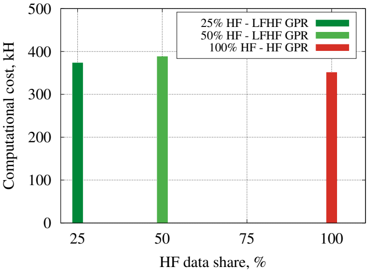




Results

HFLF GPR model

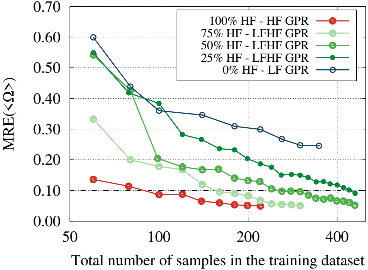


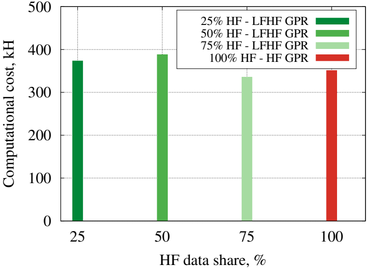


Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Results

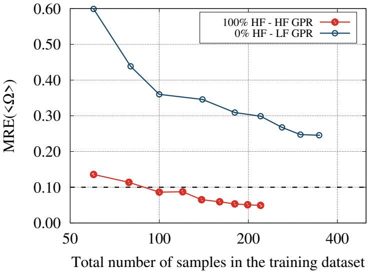
HFLF GPR model

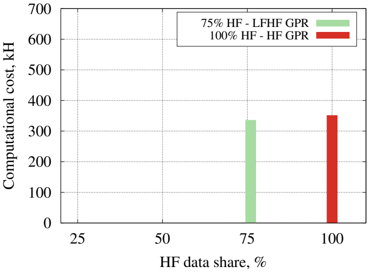




Results

LC GPR model

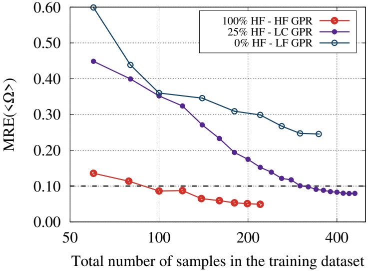


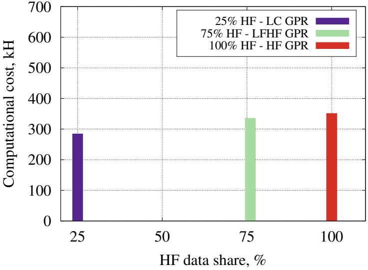


Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Results

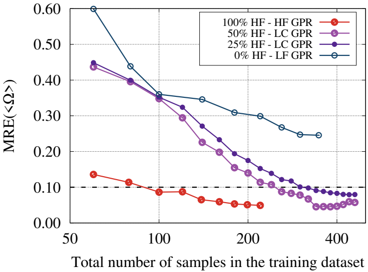
LC GPR model

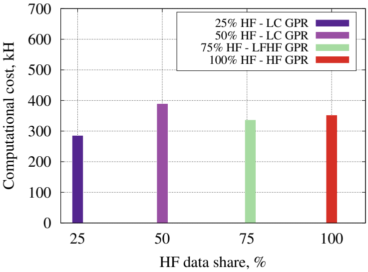




Results

LC GPR model

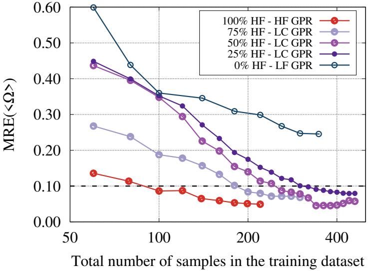


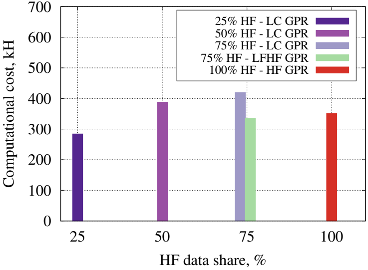


Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Results

LC GPR model

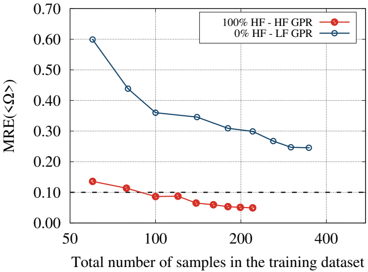


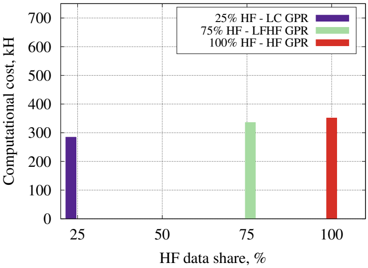


Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Results

MF GPR model

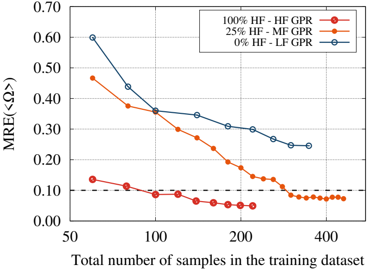


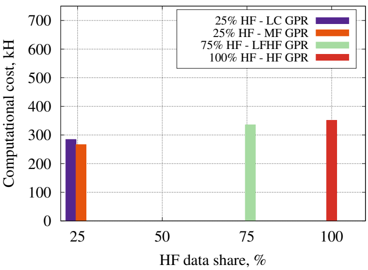


Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Results

MF GPR model

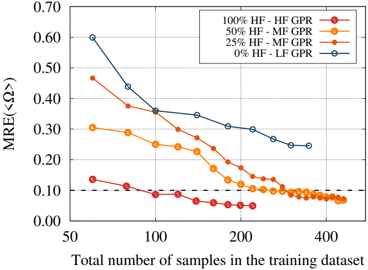


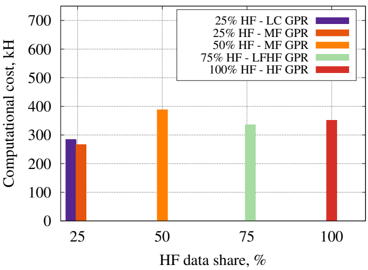


Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Results

MF GPR model

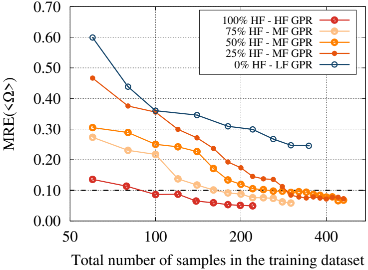


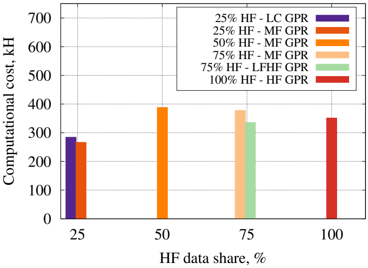


Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Results

MF GPR model

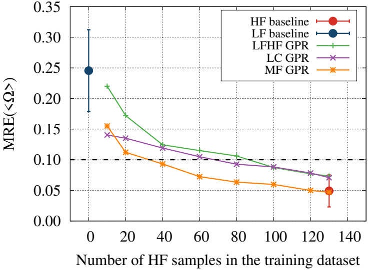


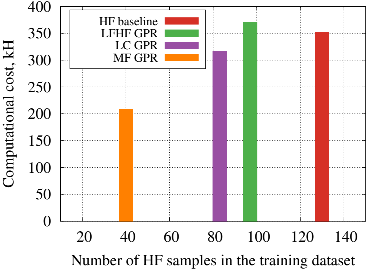


Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Results

MRE of the average enstrophy $<\Omega>$ for different number of HF training samples using different surrogate models





Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Results

Summary of the obtained results

Model	Samples			Kh	MRE	
	HF	LF	total		$<Nu>$	$<\Omega>$
HF GPR	130	-	130	351	0.046 ± 0.02	0.046 ± 0.03
LF GPR	-	350	350	100	0.193 ± 0.03	0.246 ± 0.07
HFLF GPR	120	40	160	335	0.042 ± 0.06	0.093 ± 0.05
LC GPR	75	225	240	285	0.112 ± 0.06	0.100 ± 0.05
MF GPR	40	350	380	208	0.092 ± 0.04	0.093 ± 0.05

Conclusions

Conclusions - 1

- The developed surrogate model provides almost instant accurate predictions using an ordinary office computer.
- The input data of the developed model is structured to take the values of temperature and velocity in the locations, which could be replaced by sensor readings.
- The main computational burden of the surrogate model is the cost of its development because, at this step, a comprehensive set of HF data is required.
- We spent 650Kh on 240 HF CFD simulations and 100Kh on 350 LF simulations.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Conclusions

Conclusions - 1

- The developed surrogate model provides almost instant accurate predictions using an ordinary office computer.
- The input data of the developed model is structured to take the values of temperature and velocity in the locations, which could be replaced by sensor readings.
- The main computational burden of the surrogate model is the cost of its development because, at this step, a comprehensive set of HF data is required.
- We spent 650Kh on 240 HF CFD simulations and 100Kh on 350 LF simulations.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Conclusions

Conclusions - 1

- The developed surrogate model provides almost instant accurate predictions using an ordinary office computer.
- The input data of the developed model is structured to take the values of temperature and velocity in the locations, which could be replaced by sensor readings.
- The main computational burden of the surrogate model is the cost of its development because, at this step, a comprehensive set of HF data is required.
- We spent 650Kh on 240 HF CFD simulations and 100Kh on 350 LF simulations.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Conclusions

Conclusions - 1

- The developed surrogate model provides almost instant accurate predictions using an ordinary office computer.
- The input data of the developed model is structured to take the values of temperature and velocity in the locations, which could be replaced by sensor readings.
- The main computational burden of the surrogate model is the cost of its development because, at this step, a comprehensive set of HF data is required.
- We spent 650Kh on 240 HF CFD simulations and 100Kh on 350 LF simulations.

Conclusions

Conclusions - 2

- The use of multi-fidelity models reduced the computational cost considerably.
- Even a simple HFLF GPR model is less computationally expensive than the baseline HF GPR model.
- More sophisticated multi-fidelity models like LC GPR and MF GRP required at least 1.5 times less computational resources than the HF GPR model.
- The MF GRP has shown the best trade-off between computational cost and accuracy among studied multi-fidelity models.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Conclusions

Conclusions - 2

- The use of multi-fidelity models reduced the computational cost considerably.
- Even a simple HFLF GPR model is less computationally expensive than the baseline HF GPR model.
- More sophisticated multi-fidelity models like LC GPR and MF GRP required at least 1.5 times less computational resources than the HF GPR model.
- The MF GRP has shown the best trade-off between computational cost and accuracy among studied multi-fidelity models.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Conclusions

Conclusions - 2

- The use of multi-fidelity models reduced the computational cost considerably.
- Even a simple HFLF GPR model is less computationally expensive than the baseline HF GPR model.
- More sophisticated multi-fidelity models like LC GPR and MF GRP required at least 1.5 times less computational resources than the HF GPR model.
- The MF GRP has shown the best trade-off between computational cost and accuracy among studied multi-fidelity models.

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya

Conclusions

Conclusions - 2

- The use of multi-fidelity models reduced the computational cost considerably.
- Even a simple HFLF GPR model is less computationally expensive than the baseline HF GPR model.
- More sophisticated multi-fidelity models like LC GPR and MF GRP required at least 1.5 times less computational resources than the HF GPR model.
- The MF GRP has shown the best trade-off between computational cost and accuracy among studied multi-fidelity models.

Conclusions

Future work

- Study the generalization properties of the developed methodology.
- Adapt the existing methodology to a direct prediction of comfort parameters, such as predicted mean vote.
- A more broad study on a proper choice of data fidelity is required.
- The model could be further analyzed in terms of extrapolation capabilities.

THANK YOU FOR YOUR ATTENTION!

Ready for your questions!

Centre Tecnològic de Transferència de Calor
Universitat Politècnica de Catalunya