Assessing Physics Parameterizations using Evolutionary Computation

Iciar Guerrero-Calzas[1,2][0009−0001−4385−0553]⋆, Lorenzo Rossetto[2][0009−0005−1295−528X], Ana Cortés[1][0000−0003−1697−1293], Mauricio Hanzich[2][0000−0003−4460−1118], and Josep Ramón Miró[3][0000−0003−2838−6083]

[1] Computer Architecture and Operating Systems Department. Universitat Autònoma de Barcelona. Carrer de les Sitges, s/n 08193 Cerdanyola del Vallès, Spain
iciar.guerrero@autonoma.cat, ana.cortes@uab.cat
[2] Mitiga Solutions S.L., Carrer de Julià Portet, 3, 08002 Barcelona, Spain
{iciar.guerrero, lorenzo.rossetto, mauricio.hanzich}@mitigasolutions.com
[3] Servei Meteorològic de Catalunya, Carrer del Dr. Roux, 80, 1a planta, 08017, Barcelona, Spain jr.miro@gencat.cat

Abstract. Hailstorms are intense, localized weather phenomena that can severely impact agriculture, infrastructure, and property, making precise forecasting essential for risk management. The Weather Research and Forecasting (WRF) model is widely used for numerical weather prediction, offering numerous physical parameterization options to represent atmospheric processes. However, due to the large number of possible configurations, identifying the most suitable configuration is a challenge. This research uses a genetic algorithm (GA) to systematically refine WRF physics schemes for hail prediction in Central Europe, specifically for the hail events of June 2022. Within this framework, WRF configurations are treated as individuals in a population that evolves through selection, crossover, and mutation over multiple iterations. Fitness is evaluated using the F2 score. This methodology allows to evaluate more than 2.4 million possible setups improving the WRF model's capacity to accurately represent hailstorms. This strategy provides a robust framework for testing a wide range of setups, proving its value in refining parameterizations to better forecast impactful weather phenomena.

Keywords: Hail Modeling · Genetic Algorithm · Numerical Weather Prediction (NWP)

1 Introduction

Hail is a significant weather phenomenon that poses a severe risk to agriculture, infrastructure, and property worldwide. It is a form of precipitation consisting of balls or irregular ice particles that develop within convective clouds under strong upward air currents [2]. Although relatively rare, hailstorms can cause extensive damage due to their localized and intense nature, resulting in billions

⋆ Corresponding author: iciar.guerrero@autonoma.cat

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

of US dollars (USD) in damages annually [24], making accurate prediction essential for risk mitigation and disaster preparedness. Hail formation is driven by extreme atmospheric conditions, such as strong updrafts, supercooled water, and freezing-level dynamics. These processes involve complex microphysical interactions, which make accurately predicting hail challenging using traditional numerical weather prediction (NWP) models. Accurate hail forecasting requires high-resolution simulations using advanced models like the Weather Research and Forecasting (WRF) model, which includes multiple physics parameterizations for simulating atmospheric processes such as microphysics, convection, radiation, and planetary boundary layer (PBL) interactions. Choosing the optimal combination of these schemes is critical but challenging due to the large number of possible configurations.

Research on the WRF model has highlighted the importance of optimizing physical parameterization to enhance its ability to accurately simulate severe weather events, including hailstorms. Previous studies have demonstrated the sensitivity of WRF outputs to different parameterization choices. For example, [29] showed that the local rainfall intensity during Typhoon Fitow was highly sensitive to the microphysics schemes chosen, particularly regarding graupel and hail parameterizations. Similarly, [26] highlighted the impact of different combinations of microphysics, cumulus, and PBL schemes on the forecasting of tropical cyclones, affecting their track, intensity, and rainfall. In another study [5], the authors examined the impact of land surface physics on the simulation of horizontal convective rolls in a tropical coastal environment, affecting turbulent structures and the distribution of heat. Furthermore, in [25], it was found that the choice of microphysics greatly affected updraft characteristics during severe thunderstorms in Southeast India, while [20] showed that the model performance was more sensitive to selection of microphysics schemes than to PBL options. Despite these insights, traditional sensitivity tests are limited to a subset of physics parameterization options, providing valuable but incomplete insights, as they do not explore all possible configurations and their interactions, which are crucial for accurately simulating complex phenomena like hail.

To address this challenge, we have applied evolutionary computation, in particular, a Genetic Algorithm (GA) has been used to efficiently explore a wider range of WRF physics schemes and identify optimal configurations to simulate hail events. Unlike traditional methods that test fixed or limited sets of configurations, GA applies evolutionary principles to iteratively refine WRF setups across multiple generations. GA significantly expands the solution space and systematically improves the simulation accuracy through continuous improvements [23]. In the proposed GA, the WRF configuration options are encoded as "genes" and each "individual" represents a specific configuration of the model. While GAs have successfully optimized WRF parameterizations for applications such as wind and solar energy prediction [28] and sea breeze prediction [31], their use in hail event simulations remains relatively unexplored.

This paper is organized as follows. In Section 2, the models and observational data used are described; Section 3 exposes the methodology for calibrating the

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

physics parametrizations of WRF model applying the Genetic Algorithm; Section 4 discusses the results, and Section 5 summarizes the main conclusions and future work.

2 Models and Data

2.1 Observational Data and Study Case

In this study, we have used the European Severe Weather Database (ESWD), operated by the European Severe Storms Laboratory (ESSL) [11]. This database provides detailed information on severe storm events in Europe, including thunderstorms, hail, and tornadoes. The major hailstorms of 2022 in Europe were marked by several record-breaking events, including an overall total of 8,224 large hail reports: 1,334 very large hail reports (≥ 5 cm), and 18 giant hail reports (≥ 10 cm) [12]. These storms caused significant damage and economic losses. The storms from 3 June to 5 June affected France, Germany, Belgium and parts of Italy, with hailstones up to 11 cm in diameter [12]. In particular, on June 4th, it caused widespread hail damage across central France. The estimated insured losses from these storms in France reached EUR 650-850 million, mainly due to hail, although flooding and wind also contributed [13]. This work is focused on this particular extreme event, so, the area of interest is Central Europe. Since the WRF model is a regional model, it is essential to define the simulation area, which is known as the domain. To achieve high-resolution outputs, a nesting process is required, where a high-resolution inner domain is placed within a lower-resolution outer domain. Figure 1 shows the two domains used for the studied area, where $D01$ corresponds to domain 1 with a 9 km low-resolution domain, whereas $D02$ (domain 2) is an inner-nested domain with a 3 km resolution.

2.2 The WRF-HAILCAST Model

The Weather Research and Forecasting (WRF) model is a state of the art mesoscale numerical weather prediction system. In this work, we have used the WRF-HAILCAST model, which couples WRF with the HAILCAST module [27, 1] providing a framework to simulate hailstone growth and trajectory within realistic three-dimensional atmospheric conditions rather than idealized profiles [20]. For simplicity, we will refer to it as the WRF model throughout the rest of this paper. Weather prediction involves physical processes such as evaporation, condensation, friction, radiation, and turbulent transfers of moisture, which cannot be directly resolved numerically. To emulate these processes, numerical weather prediction models like WRF use parametrizations that interact during simulations to represent atmospheric dynamics and thermodynamics. These physics schemes included in the model are summarized in Table 1.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

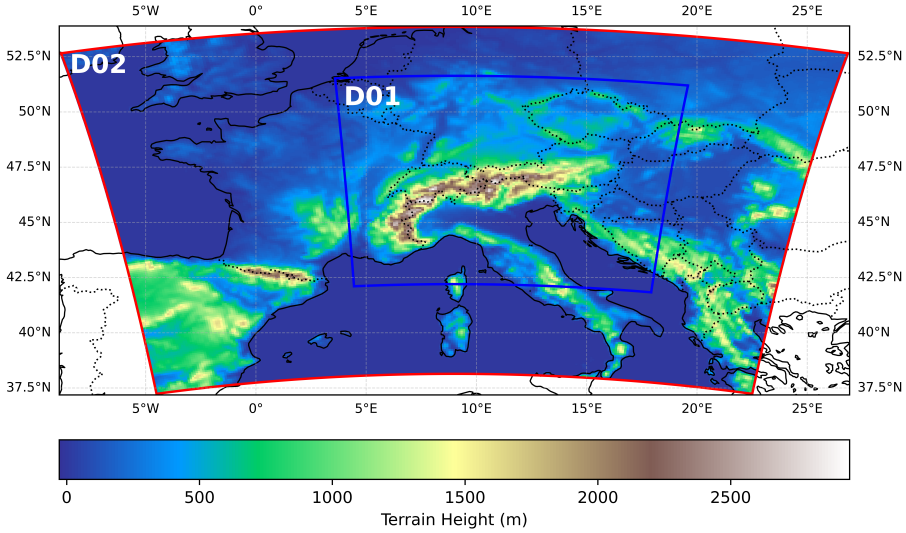

Fig. 1. WRF model domains. *D*01 9 km at resolution and *D*02 at 3 km resolution. Shading indicates topography height.

Table 1. Physics schemes included in the WRF model [21]

Physic schemes	**WRF Parameterization**	**Description**
Microphysics (MP)	mp_physics	Models cloud and precipitation processes, including rain, snow, ice, and graupel formation.
Cumulus (CU)	cu_physics	Represents the effects of convective clouds that are too small to be resolved by the model grid.
Longwave radiation (LW)	ra_lw_physics	Models the radiation emissions from atmospheric layers and surface emissivity.
Shortwave radiation (SW)	ra_sw_physics	Models solar radiation fluxes for both clear and cloudy conditions.
Planetary Boundary Layer (PBL)	bl_pbl_physics	Handles surface fluxes distribution and vertical mixing caused by boundary layer turbulence.
Surface layer (SL)	sf_sfclay_physics	Describes the exchange of energy and moisture between the surface and the atmosphere.
Land Surface Model (LSM)	sf_surface_physics	Models soil temperature, moisture, and snow water equivalent, influencing surface energy, and water fluxes.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

For this study, we have selected a subset of the available schemes based on documented incompatibilities in the WRF User's Guide [21]. These schemes include: 27 microphysics schemes, 12 cumulus schemes, 6 longwave radiation schemes, 7 shortwave radiation schemes, 10 boundary layer schemes, 3 surface layer schemes, and 6 land surface model schemes, resulting in a total of 2,449,440 possible combinations. Although the User's Guide provides detailed information about many known incompatibilities, some scheme combinations may still be incompatible but are not documented. Known incompatibilities are excluded from the initial population, as some of them are only suited for idealized tests or specific WRF compilation settings.

3 Methodology

Our methodology utilizes a GA to select a WRF model configuration, focusing on the physics schemes, to effectively simulate hail events. A Genetic Algorithm (GA) is a metaheuristic optimization method inspired by the principles of natural evolution [8]. This kind of optimization algorithm is well-suited for solving complex problems with large solution spaces by iteratively refining candidate solutions based on their performance. In this study, a GA is used to search for the WRF parameterization configuration that best reproduces the hail events in the study case. Each candidate solution, called *individual*, represents a unique combination of the seven WRF physics schemes listed in Table 1, which act as the *genes* of the individual (Figure 2).

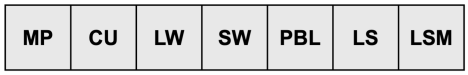

Fig. 2. Schematic representation of an individual, where each physics scheme act as "gene". Here MP, CU, LW, SW, PBL, LS, and LSM refer to microphysics, cumulus, longwave radiation, shortwave radiation, Planetary Boundary Layer, Land Surface Model, and Surface physics schemes, respectively

A set of individuals (candidate solution) defines a *population*. The GA evolves an initial population through genetic operators (selection, crossover and mutation) to identify the individual that most accurately simulates the hail events, based on a fitness score. More precisely, the GA starts by generating a random initial population of individuals. For each individual, the underlying WRF model is executed and the resulting output is compared to the observed data by evaluating a certain fitness function, which quantifies how well each individual's simulation matches the observed event. Based on these fitness scores, the next generation of individuals is created using the following operators:

– *Elitism*: The best-performing individuals from the current generation are carried over directly to the next generation without any changes. This helps

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

preserve the most successful solutions. Elitism is implemented by ranking individuals based on their fitness scores and retaining a fixed percentage of the top performers, determined by the elitism rate. In this work, 6% of the individuals have been retained as elite.
- *Selection*: Individuals are chosen for reproduction (mating population) using a linear ranking selection method. First, the population is ranked by fitness, and selection probabilities are assigned according to each individual's rank. High-ranked individuals have a greater probability of being selected for reproduction.
- *Crossover*: New individuals (offspring) are generated by mixing the genes of selected parents. Each gene of the offspring is randomly inherited from one of the parents. The crossover process is governed by a predefined crossover rate. The crossover probability in this work is set to 0.75.
- *Mutation*: Mutation adds randomness by changing some of the genes in an individual. According to the mutation rate, a subset of genes is randomly selected, and each chosen gene is replaced with a new value. In this work we have set up the mutation probability for each gene at 0.1.

These steps are iterated over n generations. After each iteration, the population is evaluated and the best individuals are carried forward. The process continues until either convergence is achieved or a maximum number of generations is reached. Convergence is determined by analyzing the fitness scores of previous generations. If the difference between the maximum fitness scores of consecutive generations is less than a specified threshold (0.001 in this case), it indicates that the algorithm is no longer making significant improvements. When this occurs on a set number of past generations, the algorithm is considered to have converged. If convergence does not occur before reaching the maximum number of generations, the process will stop regardless. Table 2 summarizes the GA configuration parameters and Figure 3 illustrates this entire GA process.

Table 2. Parameters used for the Genetic Algorithm

Parameter	
Maximum N^o of generations	35
Population size	100
Elite percentage	6%
Crossover probability	0.75
Mutation probability	0.1
Convergence threshold	0.001

3.1 Fitness Evaluation

Validating hail predictions is challenging because hail events are highly localized events, and there is a limited availability of observational data. While observational networks are essential for validation, they often fail to capture localized

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

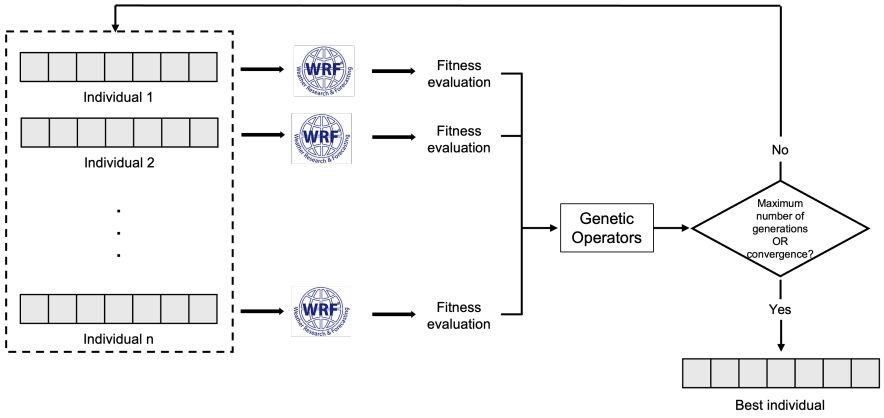

Fig. 3. Schematic illustration of the GA framework steps

hail occurrences, resulting in an over-representation of non-events. This sampling bias creates a significant class imbalance, with non-hail cases dominating the dataset. Furthermore, the rarity of hail and positional uncertainties in both forecasts and observations further complicate the validation, making traditional methods less effective. The choice of a fitness function is particularly important, as different functions can yield significantly different results, which directly influences the evaluation of hail prediction performance [7]. To evaluate each individual, the hail map is analyzed at grid cell level, where the WRF model indicates the presence or absence of hail. This assessment is a true-false event and is typically represented in a contingency table (see Table 3).

Table 3. Binary contingency table for whether or not hail is predicted.

		Observed	
		Yes	**No**
Predicted	**Yes**	Hits	False Alarms
	No	Missed	Correct Negatives

The comparison between the simulated and the real hail event presents four possible situations:

- Hits (H): hail both observed and predicted.
- False Alarms (FA): hail predicted but not observed.
- Missed (M): hail observed but not predicted.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

– Correct Negatives (CN): hail neither observed nor predicted.

As previously mentioned, the class imbalance in this problem poses a challenge for standard evaluation metrics, such as POD and FAR [4], which can be dominated by correct negatives. When a metric is influenced by non-hail events, it may produce misleadingly high scores even if it fails to detect hail. To mitigate this issue, we use the *F2* score, a member of the *F-beta* family, which ignores correct negatives and emphasizes the detection of both observed and simulated hail events (Equation 1).

$$F_2 = \frac{(1+2^2)\cdot \mathrm{H}}{(1+2^2)\cdot \mathrm{H} + 4\cdot \mathrm{M} + \mathrm{FA}} \tag{1}$$

This metric gives more weight to recall, making the metric more sensitive to false negatives, that is, we prioritize the identification of hail events, even if it leads to an increase in false alarms. This aligns with our main objective of minimizing missed detection, that is, missing actual hail events. Missing a hail event can have serious consequences, including missed warnings, insufficient preparation, and potential damage. Therefore, we prioritize ensuring that the model identifies as many hail events as possible, even if it occasionally predicts hail where it did not occur.

High-resolution weather forecasts often struggle to precisely match observed small-scale events in space, time, or intensity. These discrepancies arise due to the natural variability of the phenomena and differences between model predictions and observational data. For example, small differences in storm location may not indicate a significant forecast error but can lead to poor performance metrics when comparing the model output directly with observations at individual grid points. To address this issue, we have applied a convolutional filter that smooths the comparison between forecasts and observations, preventing small spatial misalignment from overly penalizing the model's performance. To do that, first, we upscale the point-based observational data to match the model's 3 km grid resolution, ensuring that both datasets are comparable. Once we have both maps, the observations and the predictions, at the same resolution (3 km) we overlap them to obtain the cells corresponding to hits, misses, and false alarms. These cells will be assigned a weight of 1 when considered to evaluate the *F2* score (Equation 1). Figure 4 illustrates this using a 6×6 grid map, where blue cells represent predicted hail events (False Alarms), yellow cells represent observations (Misses), and green cells highlight exact matches (Hits). The map on the left depicts this direct overlapped map. In this map, there is only one green cell corresponding to an exact hit, then, this hit will be weighted by 1 as well as the yellow and blue cells, which correspond to misses and false alarms, respectively.

Furthermore, to mitigate the impact of small spatial mismatches, the *near-hit* concept is introduced. Then, for each cell corresponding to a false alarm or a miss, the Moore neighborhood is considered (purple squares in Figure 4). The Moore neighborhood is the set of cells orthogonally or diagonally adjacent to a given cell map [32].

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

To account for spatial mismatches, we classify a false alarm (FA) or miss (M) as a near-hit if its Moore neighborhood contains: (1) if the FAs Moore neighbourhood contains any observed event (Figure 4(d)), or if the Ms Moore neighbourhood contains any predicted event (Figure 4(c)). Such cases are treated as near-hits and assigned a reduced weight of 0.5 in the F2-score calculation (Equation 1). Cases without neighbouring real events or predicted ones maintain their original classification as FAs or Ms with full weight (1.0), as illustrated in Figures 4(a-b).

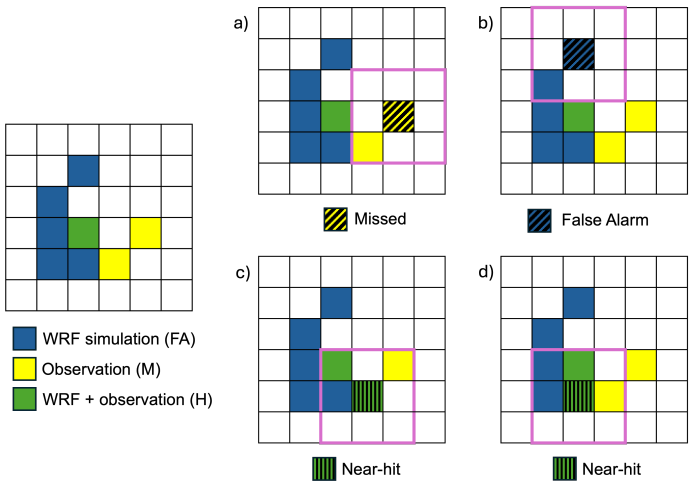

Fig. 4. Fitness evaluation using a 6 × 6 grid at 3 km resolution. Green cells are exact hits, blue cells represent WRF-predicted hail events (false alarms) and yellow cells correspond to observations (misses). Vertically striped cells are near-hits.

This approach mitigates the impact of small spatial mismatches, ensuring that they are not overly penalized as false alarms or misses. By incorporating near-hits, the validation process emphasizes the model's ability to capture the event's general location. This reduces false alarms and ensures that the model accurately predicts both the occurrence and the general location of events, even with minor spatial deviations. By accounting for spatial uncertainty, this validation method offers a fairer and more robust evaluation of high-resolution forecasts, improving their reliability in assessing hail prediction performance.

4 Experimental Results

All WRF simulations described in this section were executed on the MareNostrum 5 General Purpose Partition (GPP) supercomputer [3]. MareNostrum 5 GPP consists of 6,408 nodes powered by Intel Sapphire Rapids processors, each configured with two 8480+ CPUs running at 2GHz, providing 112 cores per node. Each standard node includes 256GB of DDR5 memory, with 216 nodes upgraded

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

to 1TB of memory and 960GB of NVMe storage for high-speed data access. The system features NDR200 interconnects, with each link shared between two nodes, delivering 100Gb/s bandwidth per node. Additionally, 72 high-bandwidth memory (HBM) nodes use Intel Sapphire Rapids 03H-LC processors, each with 112 cores at 1.9GHz and 128GB of HBM memory, offering an impressive 2TB/s memory bandwidth per node. The entire machine achieves a peak performance of 45.9 PFlops and operates on a fat-tree network topology, ensuring high-speed communication and scalability.

Our executions have been carried out on 100 nodes per generation, and a runtime limit of three hours per individual is enforced. This setup allows for the simultaneous execution of multiple WRF configurations within a generation, ensuring efficient use of resources. If a job exceeds the three-hour runtime limit, it is flagged for review. This computational setup not only enables the execution of thousands of WRF simulations within a reasonable time frame (hours), but also guarantees the reproducibility and scalability of the simulations. Using the power of MareNostrum 5's advanced infrastructure, the GA is able to thoroughly explore the complex configuration space of the WRF physics options.

The setup of the WRF model consisted of two one-way nested domains: the outer domain with a horizontal grid resolution of 9 km (260 $\times$ 190 grid points) and the inner domain with a resolution of 3 km (361 $\times$ 340 grid points) (Figure 1). Vertical resolution was maintained at 50 levels, with the model top set at 50 hPa. Boundary conditions were updated hourly using ERA5 reanalysis data with a horizontal resolution of 0.25°, 38 levels, and 1 hour temporal resolution [15]. All simulations were initialized at 05 UTC, with 6-hour spin-up, and run for a total of 40 hours to better capture the hail events and cover the development and decay of storms. Model outputs were generated hourly for both domains, with HAILCAST module activated only in the inner domain at a 60-second interval. The physics schemes changed between simulations, but certain aspects remained fixed as, for instance, the cumulus parameterization was only applied in the outer domain.

As it has been previously mentioned, although certain physics scheme combinations are known to be incompatible, there exist others that may still be incompatible but are not documented. In order not to prevent GA of exploring as much searching space as possible, when an individual is executed and it returns an incompatibility error, the fitness score assigned to the corresponding individual will be a low value (this point is explained later in this section). This way, they are considered for generating the next population but with a very low probability of being used in any genetic operator.

A total of 35 generations were run with a population size of 100 individuals, resulting in 3500 simulations. From these, 1873 unique WRF configurations were created, reflecting the extensive search space explored by the GA. This extensive search allowed the GA to identify combinations of physics parameterizations that might not have been considered through traditional sensitivity analysis methods. Although the 3500 simulations represent only a fraction of the space, they provide valuable insights into the algorithm's ability to iden-

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

tify optimal configurations and demonstrate its utility for practical applications. The best-performing configuration was identified in generation 25. The detailed breakdown of this configuration is shown in Table 4. We chose to stop at generation 35 because it allowed us to verify the stability and to ensure that the best-performing configuration remained across multiple generations.

Table 4. Best-performing WRF configuration identified by the genetic algorithm

WRF Parameterization	**Option**	**Description**
cu_physics	14	KIAPS SAS (KSAS) [14, 18]
mp_physics	14	WRF Double-Moment 5-class (WDM5) [19]
ra_lw_physics	5	New Goddard [10, 9]
ra_sw_physics	4	RRTMG [16]
bl_pbl_physics	9	UW [6]
sf_sfclay_physics	1	Revised MM5 [17]
sf_surface_physics	4	Noah-MP [22, 30]

Figure 5 shows the results of the best-performing configuration, comparing observed hail events (red dots) and WRF-simulated hail occurrences (purple) within the 3 km domain. While some discrepancies are present due to model uncertainty and observational limitations, the overall consistency reinforces the GA's ability to optimize model performance effectively. The alignment of storm clusters highlights the effectiveness of the optimized configuration, although discrepancies arise from model uncertainty and observational limitations.

The progression of the fitness score across generations is shown in Figure 6. Initially, the population consisted of random configurations, with fitness scores relatively low. The first notable improvement occurred in generation 3, where the fitness score reached 0.2273. This was followed by another increased to 0.2626 in generation 6. By generation 10 the fitness score stabilized at 0.2629. At generation 20 the fitness score improved slightly reaching 0.2635 and up to 0.2639 in generation 24. But the most significant jump occurred at generation 25, when the score reached 0.2915. This score remained steady through generation 35.

While the final fitness score of 0.2915 is not very high, it is important to consider the uncertainty of the model and the scarcity of the observational data available for validation. The observed jumps in the fitness score can be due to the chosen population size. Given the size of the search space – approximately 2.4 million possible combinations – and a relatively small population of 100 individuals, only a small fraction of the configurations is explored.

Despite these challenges, the algorithm's performance is still promising. The results from the spatial analysis (Figure 5) further support the GA's effectiveness, as the final configuration was able to capture most of the storm clusters.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

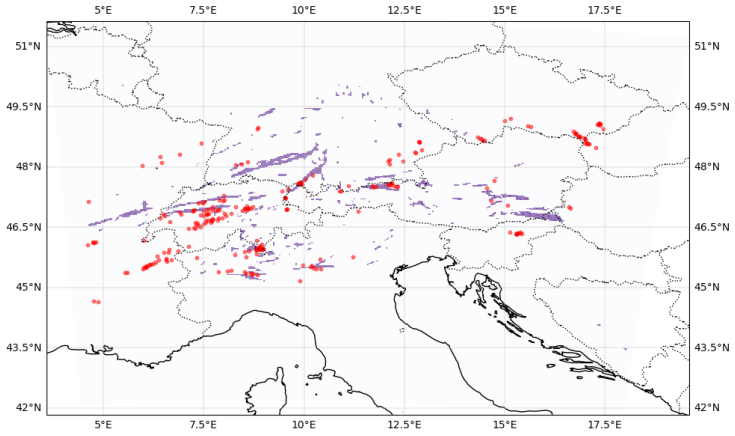

Fig. 5. Spatial comparison of observed hail events (red dots) and simulated hail occurrences (purple) using the best-performing WRF configuration. The map focuses on the inner domain, showcasing the alignment of storm clusters and the ability of the configuration to reproduce hail events.

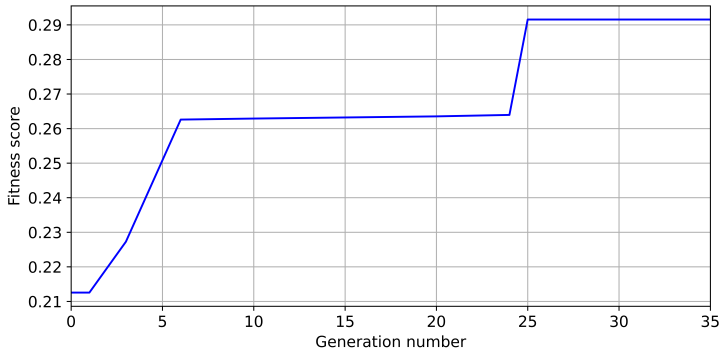

Fig. 6. Evolution of the best fitness scores across generations. Significant improvements were observed in generations 3, 6, and 25, with the score stabilizing at 0.2915 from generation 25 to 35

5 Conclusions and Further work

In this study, a Genetic Algorithm was used to assess the physics parameterizations of the WRF model for hailstorm forecasting. Despite challenges like limited observational data, model uncertainty, and the use of fixed discrete parameter choices, which limits the fine-tuning of settings, the algorithm showed promising results in identifying effective WRF configurations for hail predictions. The proposed approach has been applied in a severe hail event that took place in Central Europe on June of 2022. The results captured the majority of the ex-

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

treme storm clusters proving the ability of the system to assess discrete model configurations. Typically, there are small spatial mismatches when predicting the location of a hailstorm, thus, to mitigate this issue, the concept of near-hit has been introduced. The original 3 km map has been upscaled to a 9 km Moore neighborhood map, this way, the fitness score used in the GA process does not completely dismiss certain false alarms and misses when they are close to a exact hit but giving them partial weight in the fitness score.

Future work should consider expanding this flexibility to a larger radius, such as 25 km, to better account for the spatial characteristics of hailstorms, which often extend over 10 km-25 km due to storm motion and hailstone advection. This change would align with the needs of operational users, such as emergency managers and agricultural decision-makers, who require broader regional risk awareness for effective response. To assess the robustness of the methodology, future research should involve running the GA on multiple dates to evaluate its performance across diverse meteorological conditions. This would provide a broader understanding of the model's capabilities and its adaptability to different weather scenarios.

Acknowledgments. This work is supported by the Industrial Doctorate Plan of the Department of Research and Universities of the Generalitat de Catalunya under contract 2023DI00013. It also has been granted by the Spanish Ministry of Science and Innovation MCIN AEI/10.13039/501100011033 under contract PID2023-146193OB-I00 and by the Catalan government under grant 2021 SGR-00574. The simulations and data analysis presented in this study were supported by the Barcelona Supercomputing Center (BSC) and Red Española de Supercomputación (RES) under project Inno-2-2024-0002.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Adams-Selin, R.D., Ziegler, C.L.: Forecasting hail using a one-dimensional hail growth model within wrf. Monthly Weather Review **144**(12), 4919 – 4939 (2016). https://doi.org/10.1175/MWR-D-16-0027.1
2. American Meteorological Society: Hail. *Glossary of Meteorology* (2017), https://glossary.ametsoc.org/wiki/Hail, accessed: 2024-11-27
3. Barcelona Supercomputing Center (BSC): Marenostrum 5 supercomputer (2025), https://www.bsc.es/marenostrum/marenostrum-5, accessed: 2025-01-09
4. Bhavyasree, Panda, S., Wasson, G., Mondal, U., Kumar, A., Sharma, D.: Assessing the performance of wrf model in simulating severe hailstorm events over assam and bihar, india. Modeling Earth Systems and Environment **10**, 6013–6034 (08 2024). https://doi.org/10.1007/s40808-024-02114-z
5. Bollareddy, R., C., V.S., J R, R., Balasubramaniam, V.: Numerical simulation of horizontal convective rolls over a tropical coastal site using wrf: Sensitivity to land surface physics. Pure and Applied Geophysics **180**, 1–24 (11 2023). https://doi.org/10.1007/s00024-023-03361-4

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

6. Bretherton, C.S., Park, S.: A new moist turbulence parameterization in the community atmosphere model. Journal of Climate **22**(12), 3422 – 3448 (2009). https://doi.org/10.1175/2008JCLI2556.1
7. Carrillo, C., Artes, T., Cortés, A., Margalef, T.: Error function impact in dynamic data-driven framework applied to forest fire spread prediction. Procedia Computer Science **80**, 418–427 (12 2016). https://doi.org/10.1016/j.procs.2016.05.342
8. Chaudhry, I.A., Usman, M.: Integrated process planning and scheduling using genetic algorithms. Technical gazette **24**, 1401–1409 (10 2017). https://doi.org/10.17559/TV-20151121212910
9. Chou, M.D., Suarez, M.J.: A solar radiation parameterization for atmospheric studies. NASA Technical Memorandum NASA/TM-1999-104606/VOL15, NASA Goddard Space Flight Center (1999), 10.2172/19990060930
10. Chou, M.D., Suarez, M.J., Liang, X.Z., Yan, M.M.H., Cote, C.: A thermal infrared radiation parameterization for atmospheric studies. NASA Technical Memorandum NASA/TM-2001-104606/VOL19, NASA Goddard Space Flight Center (2001), 10.2172/20010072848
11. Dotzek, N., Groenemeijer, P., Feuerstein, B., Holzer, A.M.: Overview of essl's severe convective storms research using the european severe weather database eswd. Atmospheric Research **93**(1), 575–586 (2009), 4th European Conference on Severe Storms
12. European Storm Forecast Experiment (ESSL): Major hailstorms of 2022. *European Storm Forecast Experiment* (2022), https://www.essl.org/cms/major-hailstorms-of-2022/, last accessed: 2024-12-11
13. Guy Carpenter: Severe convective storms in europe: June 2022. *Guy Carpenter Insights* (June 2022), https://www.guycarp.com/insights/2022/06/SCS_Europe_June2022.html, last accessed: 2024-12-11
14. Han, J., Pan, H.L.: Revision of convection and vertical diffusion schemes in the ncep global forecast system. Weather and Forecasting **26**, 520–533 (08 2011). https://doi.org/10.1175/WAF-D-10-05038.1
15. Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Thépaut, J.N.: The era5 global reanalysis. Quarterly Journal of the Royal Meteorological Society **146** (06 2020). https://doi.org/10.1002/qj.3803
16. Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shephard, M.W., Clough, S.A., Collins, W.D.: Radiative forcing by long-lived greenhouse gases: Calculations with the aer radiative transfer models. Journal of Geophysical Research: Atmospheres **113**(D13) (2008). https://doi.org/10.1029/2008JD009944
17. Jiménez, P.A., Dudhia, J., González-Rouco, J.F., Navarro, J., Montávez, J.P., García-Bustamante, E.: A revised scheme for the wrf surface layer formulation. Monthly Weather Review **140**(3), 898 – 918 (2012). https://doi.org/10.1175/MWR-D-11-00056.1
18. Kwon, Y., Hong, S.Y.: A mass-flux cumulus parameterization scheme across gray-zone resolutions. Monthly Weather Review **145** (11 2016). https://doi.org/10.1175/MWR-D-16-0034.1
19. Lim, K.S., Hong, S.Y.: Development of an effective double-moment cloud microphysics scheme with prognostic cloud condensation nuclei (ccn) for weather and climate models. Monthly Weather Review **138**, 1587–1612 (05 2010). https://doi.org/10.1175/2009MWR2968.1

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8

20. Malečić, B., Telišman Prtenjak, M., Horvath, K., Jelić, D., Mikuš Jurković, P., Ćorko, K., Mahović, N.S.: Performance of hailcast and the lightning potential index in simulating hailstorms in croatia in a mesoscale model – sensitivity to the pbl and microphysics parameterization schemes. Atmospheric Research **272**, 106143 (2022). https://doi.org/10.1016/j.atmosres.2022.106143
21. National Center for Atmospheric Research (NCAR): WRF Model User's Guide. Mesoscale and Microscale Meteorology Laboratory (2023), https://www2.mmm.ucar.edu/wrf/users/wrf_users_guide, last accessed: 2024-11-11
22. Niu, G.Y., Yang, Z.L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., Xia, Y.: The community noah land surface model with multiparameterization options (noah-mp): 1. model description and evaluation with local-scale measurements. Journal of Geophysical Research: Atmospheres **116**(D12) (2011). https://doi.org/10.1029/2010JD015139
23. Oana, L., Spataru, A.: Use of genetic algorithms in numerical weather prediction. pp. 456–461 (09 2016). https://doi.org/10.1109/SYNASC.2016.075
24. Púčik, T., Castellano, C., Groenemeijer, P., Kühne, T., Rädler, A.T., Antonescu, B., Faust, E.: Large hail incidence and its economic and societal impacts across europe. Monthly Weather Review **147**(11), 3901 – 3916 (2019). https://doi.org/10.1175/MWR-D-19-0204.1
25. Rajeevan, M., Kesarkar, A., Thampi, S.B., Rao, T.N., Radhakrishna, B., Rajasekhar, M.: Sensitivity of wrf cloud microphysics to simulations of a severe thunderstorm event over southeast india. Annales Geophysicae **28**(2), 603–619 (2010). https://doi.org/10.5194/angeo-28-603-2010
26. Shenoy, M., Raju, P., Prasad, V., Prasad, K.: Sensitivity of physical schemes on simulation of severe cyclones over bay of bengal using wrf-arw model. Theoretical and Applied Climatology **149**, 1–15 (05 2022). https://doi.org/10.1007/s00704-022-04102-8
27. Skamarock, C., Klemp, B., Dudhia, J., Gill, O., Liu, Z., Berner, J., Wang, W., Powers, G., Duda, G., Barker, D.M., Huang, X.: A description of the advanced research wrf model version 4 (2019), https://api.semanticscholar.org/CorpusID:196211930
28. Sward, J., Ault, T., Zhang, K.: Genetic algorithm selection of the weather research and forecasting model physics to support wind and solar energy. SSRN Electronic Journal (01 2022). https://doi.org/10.2139/ssrn.3999150
29. Xu, H., Zhang, D., Li, X.: The impacts of microphysics and terminal velocities of graupel/hail on the rainfall of typhoon fitow (2013) as seen from the wrf model simulations with several microphysics schemes. Journal of Geophysical Research: Atmospheres **126**(23) (2021). https://doi.org/10.1029/2020JD033940
30. Yang, Z.L., Niu, G.Y., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., Tewari, M., Xia, Y.: The community noah land surface model with multiparameterization options (noah-mp): 2. evaluation over global river basins. Journal of Geophysical Research: Atmospheres **116**(D12) (2011). https://doi.org/10.1029/2010JD015140
31. Yoon, J.W., Lim, S., Park, S.K.: Combinational optimization of the wrf physical parameterization schemes to improve numerical sea breeze prediction using micro-genetic algorithm. Applied Sciences **11**(23) (2021). https://doi.org/10.3390/app112311221
32. Zaitsev, D.A.: A generalized neighborhood for cellular automata. Theoretical Computer Science **666**, 21–35 (2017). https://doi.org/10.1016/j.tcs.2016.11.002

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97629-2_8