



Structural Equation Modelling of Detector Response in Hadron-Initiated Extensive Air Showers: A PLS-SEM Approach

Vidika Mirpuri

ABSTRACT

The project aims to investigate concepts in detector response in the extensive air showers area by means of structural equation modelling. Three concepts were explored: muon detection by absorption in the soil, radiofrequency (RF) antenna detection, and hybrid measurements which are made using the electron-photon and hadronic components of the shower. The ideas were utilized in the form of R1, R2 and R3 i.e. as latent variables and were examined through 15 factors of Likert scale and Partial Least Squares Structural Equation Modeling (PLS-SEM). There were 98 entries in the records and were processed using Smart PLS 4. While the Cronbach's alpha was at 0.867 for R1, it was found to be at 0.838 for R2 and at 0.815 for R3. As for composite reliability (ρ_c), the figures for R1, R2 and R3 were 0.904, 0.882 and 0.865 respectively. Average variance extracted (AVE) was at 0.652, 0.600 and at 0.521 respectively. Reliability and convergent validity were acquired by all three components because of these statistics. The design involved two paths which were theorized based on concepts $R1 \rightarrow R2$ and $R1 \rightarrow R3$ and had standardized coefficients of path 0.360 and 0.405 respectively. In general, the study demonstrated how physics problem can be reformulated into a range of measurable latent concepts that are assessed using PLS-SEM.

Index Terms-Extensive air showers, detector response modelling, hadronic interaction models, radio detection, gamma-hadron separation, PLS-SEM, SmartPLS.

I. INTRODUCTION

When an energetic hadron strikes the Earth's atmosphere the result is a particle shower. A primary hadron at above 10^{18} eV is observed through the secondary particles created, muon production and the shower maximum height ($X_{\max}$), which has raised some questions. In light of various models that estimate the hadronic process, the precise connection between primary particle, muon production and specific shower maximum height ($X_{\max}$) is challenging due to detector responses and nonlinearity. The concept presented in this article is useful in understanding the principle of the Pomeranchuk-Migdal (LPM) effect and determining new physics and or exotic types of primary particles under reach conditions of energy [1].

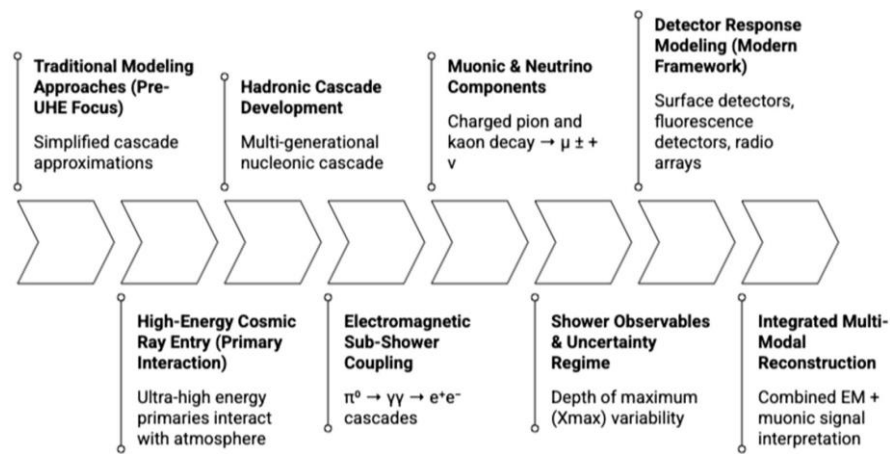


Fig. 1. Chronological Evolution of Cosmic Ray modelling

The improvement in air shower reconstruction is a result of advances in cascade physics, particle interaction, shower development as well as detector performance in cosmic ray modeling. This makes it necessary to understand the physics underlying the air shower phenomenon. Extensive air showers, which occur in the atmosphere, are initiated by a series of collisions between primary cosmic ray hadrons (which have energy above 10^{13} eV). This gives rise to a hadronic cascade with a strong longitudinal component, which is also characteristic of air showers. The interactions in the atmosphere lead to the production of large amounts of μ and neutrinos and an electromagnetic component as a result of neutral pion decay [2]. These shower properties can be reconstructed using radio-emission measurements. This model for asymmetric radio-emission footprints utilizes d_{max} and E_{geo} observables to achieve 5% energy resolution for inclined air showers up to 10-20 eV. By parameterising geomagnetic and charge-excess interference, the framework enhances primary mass discrimination and provides an efficient solution for large-scale detector response modelling [3]. Shower morphology also plays a key role in detector response and event classification. Electromagnetic subshowers produced by hadronic cascades pose significant background for IACTS. However, γ -ray showers are tightly beamed due to Coulomb scattering, whereas hadronic showers exhibit larger lateral extension and greater fluctuations from transverse momentum. Variations in the depth of shower maximum across primary masses provide critical morphological signatures for detector response modelling [4].

Modelling these signatures relies on accurate hadronic interaction models. This study evaluates hadronic interaction models, EPOS-LHC, QGSJET-II-04, and Sibyll 2.3d, by fitting two-dimensional distributions from Pierre Auger Observatory hybrid data. Current model predictions, with major experimental and simulation discrepancies, require shifts toward deeper X_{Max} and enhanced hadronic ground signals to improve data description with a statistical significance exceeding 5σ [5]. Current model uncertainties persist in the 0.1–100 TeV regime, critical for gamma ray observations. This study evaluates model predictions of air shower parameters and initial interaction microphysics that drive macroscopic shower behaviour. Substantial discrepancies in ground particle densities emerge at 100 GeV, due to divergent model treatments of early stage shower development [6]. These uncertainties motivate dedicated detector developments. The TAIGA muon scintillation stations were introduced to enhance TeV-PeV gamma ray astronomy. Utilizing CORSIKA and GEANT4 simulations for primary energies of 100–3000 TeV, the study optimizes trigger efficiency and determines an optimal 2 m soil absorber depth for hadronic-to-electromagnetic shower suppression. Results detail identification efficiencies and suppression factors critical for robust gamma-hadron separation [7].

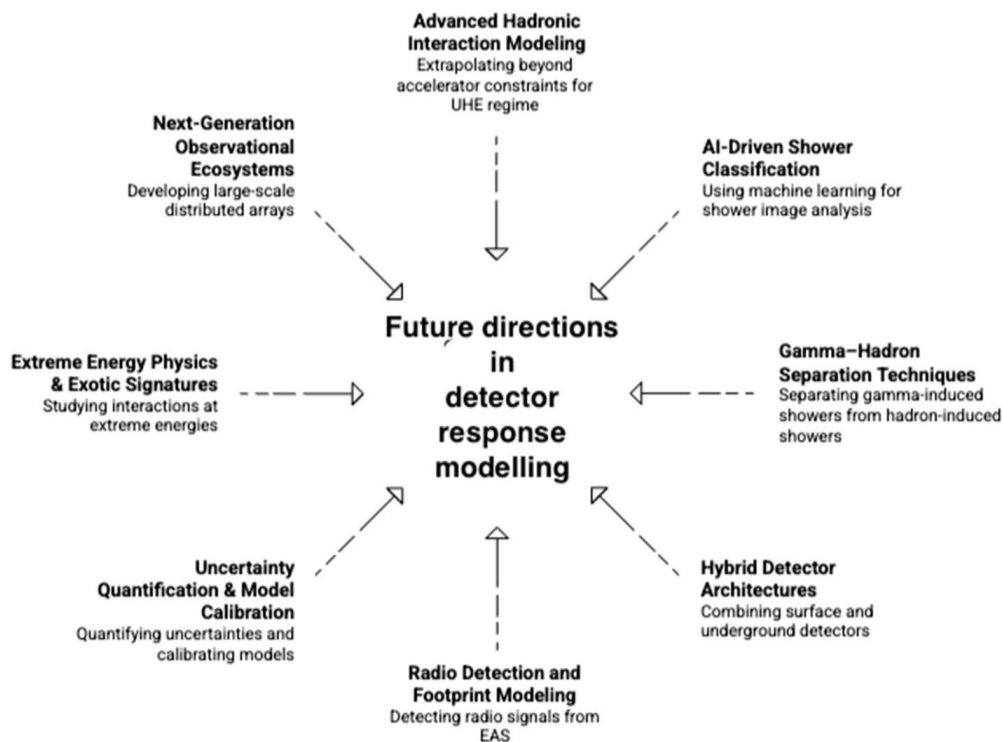


Fig. 2. Future directions in Hadron-Initiated Extensive Air Showers and Detector Response Modelling

Apart from detector enhancement, significant emphasis is being given to the creation of scalable detection systems. The system in use in the CREDO (Cosmic Ray Extremely Distributed Observatory) project is capable of detecting Cosmic Ray Ensembles by studying the detection probabilities according to parameters such as particle density, energy, mass and azimuth angle. The detector is able to function even with high energies, meaning that there will be possibilities for further improvements in its design and discrimination abilities of photons and hadrons [8]. Thanks to such achievements, contemporary hybrid detector systems have been established. Additionally, the development of cost-effective antenna technology has brought back the use of radio detection methods for large air showers that enable observing cosmic ray energies and masses effectively with a 95% operating time better than conventional Cherenkov and fluorescence methods. However, its application requires radio-quiet environments to overcome low signal-to-noise ratios relative to anthropogenic electromagnetic interference [9]. Antenna arrays rival traditional techniques as a precise, economical standard for hadronic cascades in measuring arrival direction, energy and $X_{\max}$. Hybrid systems that combine radio and muon detectors improve mass composition and gamma-hadron separation to a great extent. This leads to a more precise detection of PeV and EeV photons. The ability to scale radio techniques plays an important role in new-generation astronomy [10]. The Astroneu Array consists of independent stations with 9 scintillators and 3 radio frequency antennas used for the detection of Extensive Air Showers. Additionally, the system's hardware architecture includes remote DAQ software and offline simulation programs. The initial operational data of the hybrid supports the hypotheses related to its detection efficiency and the value of scintillator technologies [11].

Partial Least Squares Structural Equation Modelling (PLS-SEM) is used for the purpose of distinguishing true shower properties from instrumental influences in this study. This is done through the use of a novel flexible air shower generator and an effective electromagnetic simulation algorithm, which allows to explore atmosphere cascades above the GZK cut-off, from the UHECRs and to reveal important limitations in the ability to separate the primary composition, with the help of the understanding of scaling violations in the fragmentation region and the limits of the hadronic interactions uncertainties. The aim of this research is to compare the synchronous detection of high-energy showers at distant stations with particular emphasis on precise directional reconstruction. Experimental results are compared to detailed Monte Carlo simulations to assess the system's

effectiveness for using both low threshold TeV observations and ultra-high energy RF measurements. Machine learning can provide additional benefits [12] other than hardware enhancements. As part of this effort, we built neural network models to enhance separation of gamma-hadron events in pre-2014 Tibet ASy data. The Convolutional Neural Network (CNN) gave better results compared with Multilayer Perceptrons (MLPs) with AUCs of 0.781 at 10 TeV and 0.901 at 100 TeV. When applied to Crab Nebula simulations, it boosted the significance of gamma-ray detection by 1.3–1.8×, showcasing the ability of deep learning to be a promising alternative to muon discrimination using hardware-based methods in legacy air shower arrays [13]. A hybrid detector facility is installed at NEVOD Experimental Complex, which allows to study the electron-photon component and hadronic component of extensive air showers simultaneously. The system uses specialized scintillation layers and detailed simulations to assess neutron-flux measurements and detector configurations over a range of energies for the purpose of high-fidelity data for models of how detectors respond to cosmic-rays [14].

II. LITERATURE REVIEW

There is a great number of papers that reveal that translations of collider-constrained hadronic interaction models (e.g. QGSJET, SIBYLL, and EPOS) into ultra-high energy range are limited, especially in relation to the “muon problem”. Nevertheless, a simple model can change ordinary event generators allowing to see the influence of some parameters on EAS observations which provides important information about the cascade, ignoring the model itself, which can help in understanding the data with the help of Heitler's cascade [15]. The uncertainty of hadronic interaction models takes place due to an insufficient number of theoretical studies on soft hadronic and nuclear processes within the cosmic ray domain. When extrapolating to 10^{20} eV the different interaction parameters were studied carefully. Existing literature shows that the changes in them have a direct influence on inelastic cross sections and the generation of secondary particles between 10^{12} and 10^{19} eV, thus complicating the EAS comprehension [16]. Using temporal feature extraction techniques as well as sophisticated ensemble methods such as gradient boosting and stacking is the way to improve the discrimination and decrease such an error as a systematic one. Transitioning from standard Hillas parameterization to deep learning architectures like Graph Neural Networks is essential in improving classification and regression accuracy. These techniques form the foundation of EAS response modelling [17].

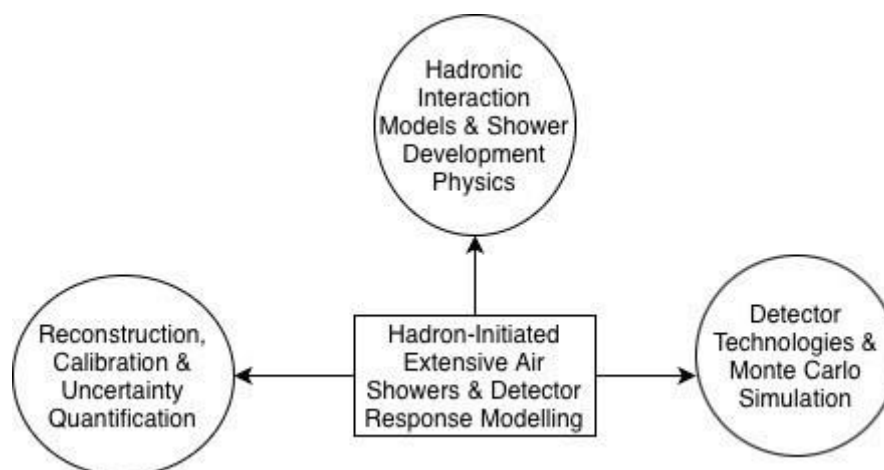


Fig. 3. Hadron-Initiated Extensive Air Showers and Detector Response Modelling

Despite these computational advances, accurate detector modelling ultimately depends on reliable descriptions of hadronic shower development. Characterization of cosmic rays above 10^{15} eV necessitates indirect detection via EAS, requiring precise simulations for primary particle reconstruction. Particularly, modelling hadronic multiparticle production at energies up to 10^{20} eV, significantly exceeding the standard reach. Studies reviewing the predictive uncertainties of interaction models, emphasise the bidirectional relationship between accelerator-based measurements and air

shower observations for advancing particle physics [18]. The IACT technique detects particle induced atmospheric radiation to enable observation of VHE gamma rays. Stereoscopic arrays enhance angular and energy reconstruction but the overwhelming hadronic cosmic ray background persists. However, leveraging morphological diversity and machine learning reduces background effects. Additionally, hadronic showers exhibit distinct temporal evolutions as established in IACT analysis, but their application within the ASTRI mini array framework is underexplored. This study characterises several temporal metrics, evaluating their effectiveness in improving background rejection compared to standard spatial reconstructions [19].

To reduce model dependence, researchers have also explored universal properties of shower development. Research on shower universality identifies model independent scaling in longitudinal profiles and particle densities, enabling composition sensitive analyses via X_{max} . Monte Carlo frameworks like CORSIKA measure the sensitivity of cascade development to primary mass and late stage interactions. They utilise specialized hadronic and electromagnetic packages, at a standard 80 GeV threshold. While Model induced variations tend to diminish at energies approaching 100 TeV, ground level data from early cascade development is leveraged. Hence, comparing these early stage dynamics is essential for advancing accuracy of detector response modelling [20]. Standard High energy particle showers exhibit universality across media including air, water and ice. Parametrized approximations are used to model Cherenkov light yield and ameliorate the complexities of Monte Carlo simulations. Constructing distributions for parameters yields intricate event-to-event descriptions of fluctuations in amplitude and longitudinal profile. These stochastic effects are instrumental for signal and background characterisation [21]. Monte Carlo radiation modelling also reduces reliability risks faced by exascale computing. In other words, it is possible to estimate failure-in-time rates and single-event upsets by characterizing site-specific barometric coefficients and seasonal variations. This allows for proactive allocation of resources and prevention of cosmic transients in computing environments. Therefore, radiation-induced hardware failures can be anticipated [22].

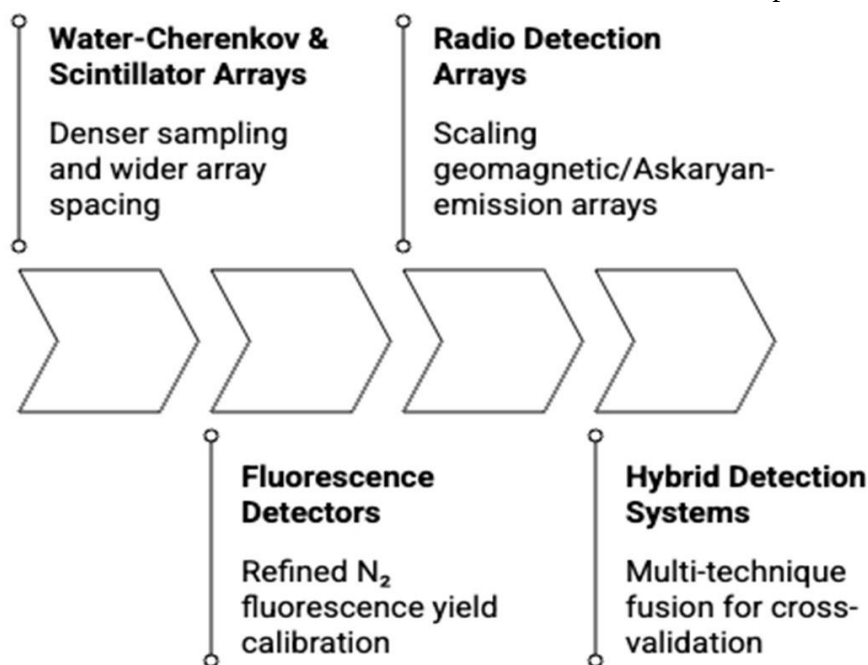


Fig. 4. Trends and Challenges in EAS observation Technologies

These modelling techniques are now being combined with increasingly sophisticated reconstruction techniques in current observational systems. In Cerenkov observations, carried out by the H.E.S.S. experiments, model driven fit algorithms are used to determine the primary particle trajectory and shower properties. These parameters assume that the resolution in direction for energies above 1 TeV is less than 0.1° . This is a much more accurate method than the traditional methods below 5 TeV, improves gamma/hadron separation and is essential for the achievement of arc-second scale source

detection [23]. The simulations suggest that the local age parameter of EAS is a useful discriminator for separating gamma and hadrons for multi-Tev energies. In experiments without any facilities for the measurement of muons, the particle identification can be improved by assigning a mean LAP to each event. This metric is a good alternative background rejection measurement particularly in instrumentation environment without classification hardware [24]. Other primary cosmic-ray compositions can be also explored using these methods. A critical benchmarking of specific hadronic species (He, Li, C and Fe) via energy estimation in Monte Carlo simulations allows a critical comparison with existing proton initiated models. This comparative study provides a deeper insight into shower development and energy scaling for different primary shower compositions. These results can serve as fundamental inputs to the structural modelling of the response of a detector to various hadronic intensities [25].

III. RESEARCH METHODOLOGY

High-energy cosmic rays, upon interacting with the atmosphere, produce extensive air showers containing both electromagnetic and hadronic components. Detector systems are built to measure these components and, from them, infer properties of the primary cosmic ray. The project is structured around three research questions: whether soil absorption can improve suppression and separation in muon detectors; whether RF antennas offer operational or measurement advantages, and whether hybrid measurements can enhance detector-response modelling [26]. PLS-SEM is particularly useful in this context, since a concept such as detector efficiency cannot be captured through a single observation. Instead, a set of observable questionnaire statements can be treated collectively as indicators of an underlying latent construct [27].

A. Research Questions and Hypotheses

TABLE I

RESEARCH QUESTIONS, HYPOTHESES, AND LATENT CONSTRUCTS

ID	Research question	Hypothesis
RQ1->R1	Does a 2 m soil absorber improve hadron-to-electromagnetic shower suppression in TAIGA muon detectors?	H1: A 2 m absorber improves shower suppression and detection efficiency.
RQ2->R2	Is RF antenna detection more efficient in duty cycle and cost without losing sensitivity to cosmic-ray energy and mass?	H2: RF antenna detection achieves higher duty cycle with comparable sensitivity.
RQ3-R3	Does combining electron-photon and hadronic measurements improve detector-response modelling accuracy?	H3: Combined measurements improve model fidelity over single-component data.

Table I links each physics question to one latent construct. R1 represents soil absorber/shower-suppression concepts, R2 represents RF-antenna performance, and R3 represents hybrid detector response. These constructs form the measurement side of the PLS-SEM framework.

B. PLS-SEM Framework

The PLS-SEM framework operates on two levels. The first is the measurement model, in which each latent construct is measured through five observed items: R1 is represented by items Q1 through Q5, R2 by Q6 through Q10, and R3 by Q11 through Q15. The second is the structural model, where SmartPLS estimates the relationships that exist among the latent constructs themselves. Within the SmartPLS model used here, R1 is specified as having direct paths to both R2 and R3 [28].

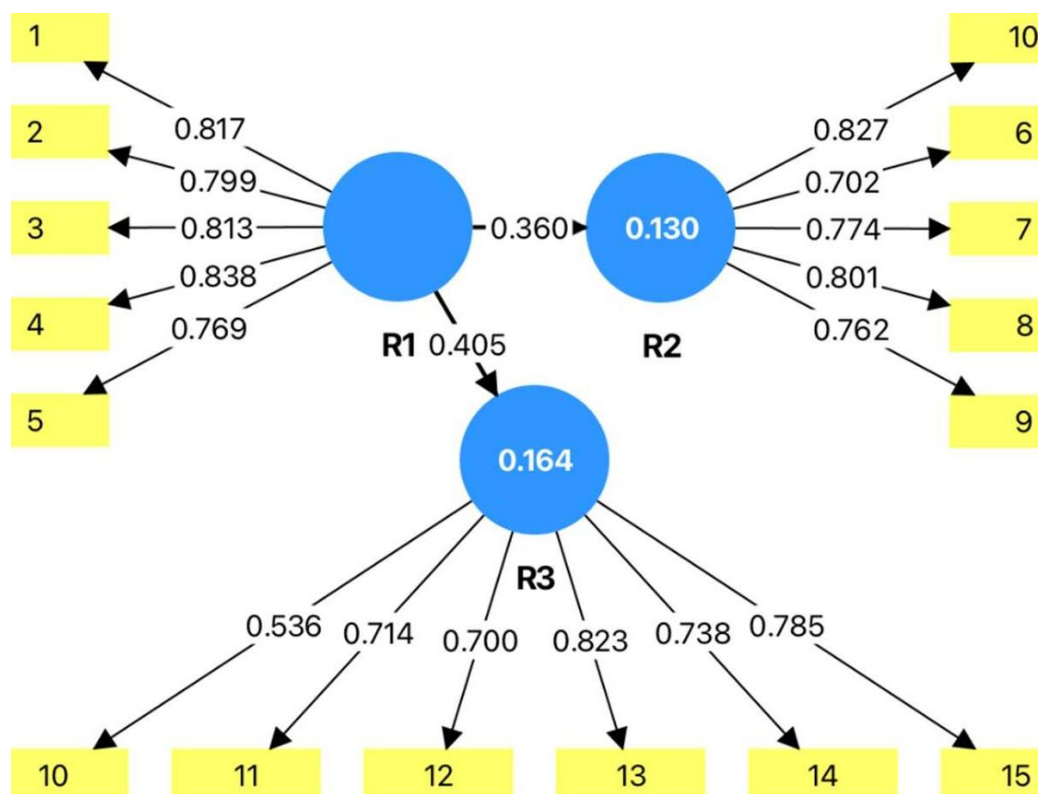


Fig. 5. PLS-SEM framework used in the uploaded SmartPLS analysis.

The circles/boxes represent latent constructs, and the arrows represent hypothesized structural relationships. The coefficient $\beta = 0.360$ means that, in the standardized model, a one-standard-deviation increase in R1 is associated with an estimated 0.360-standard-deviation increase in R2. Similarly, $\beta = 0.405$ represents the R1→R3 relationship. The five observed questionnaire items are attached to each latent construct in the SmartPLS measurement model. Importantly, this is the structural model contained in the uploaded SmartPLS report; it does not by itself test H1 as a path coefficient because R1 is an exogenous construct. If the final paper intends H1-H3 to be tested as causal hypotheses, the structural model should be redesigned so that each hypothesis corresponds to an explicit path and an appropriate endogenous outcome.

The study uses a quantitative, exploratory PLS-SEM design. The questionnaire contains 15 statements measured on a five-point Likert scale from 1 = Strongly Disagree to 5 = Strongly Agree [26]. Five items are assigned to each latent construct. For this student demonstration, 98 observations were generated to create internally coherent item responses. The dataset is therefore suitable for testing the statistical workflow but must not be represented as real participant data.

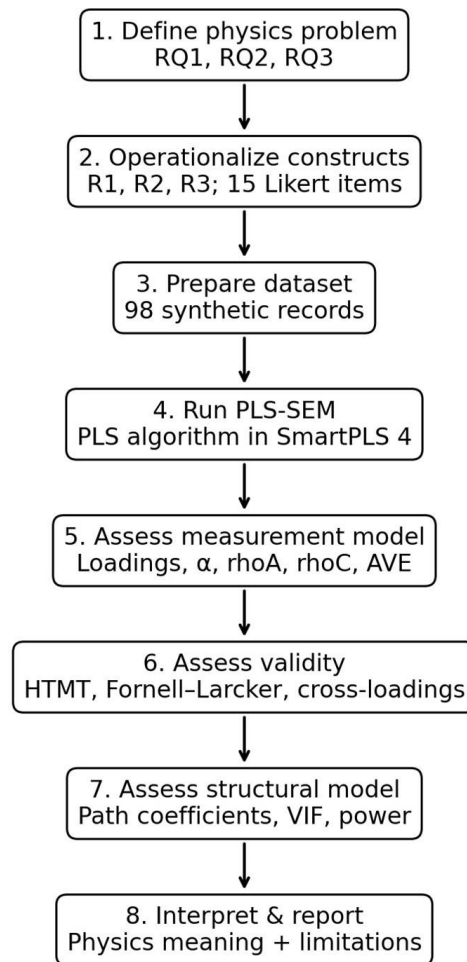


Fig. 6. Complete research and PLS-SEM analysis workflow.

The workflow runs from the initial physics questions through to the final interpretation. The measurement model is checked for adequacy before any structural relationships are interpreted [29]. The order matters because poorly constructed indicators can obscure the structural results that come after them.

TABLE II

MEASUREMENT-MODEL EVALUATION CRITERIA

Measurement-Model Criteria

Statistic	Purpose	Common interpretation
Outer loading represents its construct	Shows how strongly an item	Preferably ≥ 0.70 ; lower values may require theoretical review
Cronbach's alpha	Tests internal consistency	≈ 0.70 or higher is commonly acceptable
rho_A / rho_C	PLS-SEM reliability measures	≈ 0.70 or higher is commonly acceptable
AVE	Measures convergent validity	≥ 0.50 is commonly used
HTMT 0.85 support distinct constructs	Tests discriminant validity	Values below common cut-offs such as
VIF <5 is a common practical criterion	Checks collinearity	Lower values indicate less collinearity;

Table II gives the decision rules used to interpret the SmartPLS output [27], [29]. These are guidelines rather than automatic pass/fail rules. A student researcher should also consider the wording and theoretical meaning of each item before deleting any indicator.

IV. RESULTS

A. Outer Loadings

TABLE III
SMARTPLS OUTER LOADINGS FOR THE 15 INDICATORS

Construct	Item	Outer loading	Assessment
R1	Q1	0.817	Strong
R1	Q2	0.799	Strong
R1	Q3	0.813	Strong
R1	Q4	0.838	Strong
R1	Q5	0.769	Strong
R2	Q6	0.702	Strong
R2	Q7	0.774	Strong
R2	Q8	0.801	Strong
R2	Q9	0.762	Strong
R2	Q10	0.827	Strong
R3	Q11	0.714	Strong
R3	Q12	0.700	Below 0.70; review
R3	Q13	0.823	Strong
R3	Q14	0.738	Strong
R3	Q15	0.785	Strong

Across R1 and R2, all indicator loadings exceed 0.70. Within R3, every indicator likewise exceeds this value with the exception of Q12, which sits precisely at the conventional threshold of 0.70 [27]. This pattern suggests that, on the whole, the indicators serve as reasonably strong representations of their assigned constructs within the uploaded model.

B. Reliability and Convergent Validity

TABLE IV
CONSTRUCT RELIABILITY AND CONVERGENT VALIDITY

Construct	α	rho_A	rho_C	AVE	Decision
R1	0.867	0.875	0.904	0.652	Acceptable
R2	0.838	0.876	0.882	0.600	Acceptable
R3	0.815	0.813	0.865	0.521	Acceptable

For R1, the observed values were $\alpha = 0.867$, rho_A = 0.875, rho_C = 0.904, and AVE = 0.652. R2 showed $\alpha = 0.838$, rho_A = 0.876, rho_C = 0.882, and AVE = 0.600, while R3 produced $\alpha = 0.815$, rho_A = 0.813, rho_C = 0.865, and AVE = 0.521. On this basis, all three constructs satisfy the commonly used reliability thresholds [30], and since each AVE value exceeds 0.50, convergent validity is supported by the data at hand [31].

C. Discriminant Validity - HTMT

TABLE V
HETEROTRAIT-MONOTRAIT RATIO (HTMT) RESULTS

Construct pair	HTMT
R1 ↔ R2	0.385
R1 ↔ R3	0.448
R2 ↔ R3	0.571

The HTMT ratios came to 0.385 for the R1-R2 pair, 0.448 for R1-R3, and 0.571 for R2-R3. As all three values fall below the commonly applied thresholds [32], the constructs can be regarded as sufficiently distinct from one another within the simulated model.

D. Discriminant Validity - Fornell-Larcker

TABLE VI
FORNELL-LARCKER DISCRIMINANT-VALIDITY MATRIX

Construct	R1	R2	R3
R1	0.808	-	-
R2	0.360	0.774	-
R3	0.405	0.556	0.721

The diagonal entries correspond to the square roots of the AVE values: 0.808 for R1, 0.774 for R2, and 0.722 for R3. In each case, the diagonal value exceeds the correlations between that construct and the others appearing in the same row or column, which supports discriminant validity as assessed by the Fornell-Larcker criterion [31].

E. Cross-Loadings

TABLE VII
CROSS-LOADINGS OF INDICATORS ON THE THREE
CONSTRUCTS

Indicator	R1	R2	R3
1	0.817	0.309	0.270
Indicator	R1	R2	R3
10	0.378	0.827	0.536
11	0.217	0.260	0.714
12	0.218	0.232	0.700
13	0.305	0.296	0.823
14	0.235	0.168	0.738
15	0.294	0.336	0.785
2	0.799	0.227	0.291
3	0.813	0.272	0.326
4	0.838	0.353	0.389
5	0.769	0.272	0.341
6	0.142	0.702	0.291
7	0.226	0.774	0.429
8	0.293	0.801	0.386
9	0.258	0.762	0.437

Each questionnaire item loads most strongly on the construct it was designed to measure: Q1-Q5 load most heavily on R1, Q6-Q10 on R2, and Q11-Q15 on R3. This pattern supports the intended allocation of items to their respective constructs and offers a further check on discriminant validity [32].

F. Collinearity and Structural Results

TABLE VIII
STANDARDIZED STRUCTURAL PATH COEFFICIENTS

Structural path	Path coefficient β	Interpretation
R1 $\rightarrow$ R2	0.360	Positive relationship
R1 $\rightarrow$ R3	0.405	Positive relationship

The estimated path coefficient for R1 to R2 is $\beta = 0.360$, and for R1 to R3 it is $\beta = 0.405$. As both coefficients are positive, the model indicates that the higher the R1 score the higher are the scores on R2 and R3. However, the final results section of the report uploaded does not provide any p-values or confidence intervals, and so the coefficients cannot be used to determine if the relationships are statistically significant [27].

G. Model Fit

TABLE IX SMARTPLS MODEL-FIT SUMMARY

Model	SRMR	Interpretation
Saturated model	0.127	Above the often-used 0.08 guideline; interpret cautiously
Estimated model	0.147	Above the often-used 0.08 guideline; interpret cautiously

The report uploaded includes SRMR values of 0.127 (saturated model) and 0.147 (estimated model). Both are higher than the generally reported threshold of 0.08 [33]. The reliability and validity results are satisfactory, however the overall model fit should be interpreted with some caution. This is especially important since only the fit indices of the PLS-SEM should not be used alone to make conclusions.

H. Power and Sample-Size Diagnostics

TABLE X
POST-HOC POWER AND REQUIRED SAMPLE-SIZE DIAGNOSTICS

Path	Power $\alpha=1\%$	Power $\alpha=5\%$	N	min: N	min: N	min: N	min:
			1%,80%	5%,80%	1%,90%	5%,90%	
R1 → R2	0.892	0.972	78	48	101		67
R1 → R3	0.954	0.991	62	38	80	53	

For the R1→R2 path, the reported achieved power is 0.892 at $\alpha = 1\%$ and 0.972 at $\alpha = 5\%$. For R1→R3, it is 0.954 and 0.991, respectively. The report also estimates minimum sample sizes under different significance and power assumptions [34].

V. DISCUSSION

The results of the measurement model, which were uploaded, reveal a significantly better measurement model than the one from the previous data set. The internal consistency of all three constructs is acceptable (all alpha values are greater than 0.80) and all AVEs are greater than 0.50. The outer loadings also demonstrate the best fit of each indicator to its construct. Combined this allows the assumption that R1, R2 and R3 represent three latent variables in the simulated measurement model. R1 has the best reliability and convergent validity with the AVE of 0.652. R2 does well too, with an alpha of 0.838 and an AVE of 0.600. Comparably, the smallest AVE for the three is 0.521 for R3, which is still above the common threshold of 0.50. This is supported by the further results that were obtained with the HTMT and Fornell-Larcker results, as they demonstrated that the constructs were not the same construct that were just different labels for the same concept. The two paths, R1→R2 and R1→R3 are both pathways with positive values. However, these coefficients must be used with caution as a coefficient alone does not prove causality or significant. The uploaded report does not provide t-values, p-values or confidence interval values that are bootstrapped, so the null and alternative hypotheses cannot be formally rejected or accepted. This paper therefore merely does a descriptive report of the path coefficients and suggests bootstrapping should be done prior to any inferential statements. It also

is the case that there is a gap between the hypotheses and the model. The original hypotheses state three separate physics-based claims; however, the SmartPLS output only shows the structural model with $R1 \rightarrow R2$ and $R1 \rightarrow R3$. The model, as it is now, assumes that R1 is an explanatory model that influences R2 and R3. The student will either rewrite the hypotheses in this form or develop a structural model with paths that match the hypotheses. Both are better science than fiddling with the data to arrive at the statistical results desired.

VI. CONCLUSION & FUTURE SCOPE

The PLS-SEM analysis provides an easy-to-follow example of the translation of detector-research concepts into a measurable statistical model. All R1, R2, and R3 are acceptable with regard to reliability and convergent validity, and the results of the HTMT and Fornell-Larcker indicate that discriminant validity is supported. The model also illustrates positive standardised relationships between R1 and R2 as well as R1 and R3. This means that PLS-SEM is a pedagogically relevant tool to bridge the gap between physics, questionnaire design and multivariate statistics. However, there are a few drawbacks to consider: The questionnaire is not a direct measurement of detector physics, but is focused on perceptions and conceptual judgements. The structural model contains only $R1 \rightarrow R2$ and $R1 \rightarrow R3$ pathways, which are not necessarily one-to-one with the original H1-H3 pathways. No significance statistics is displayed (i.e. no p-value and no formal hypothesis acceptance) and can not be claimed from the current tables. Overall model fit should be interpreted with caution because the SRMR values are higher than the typical threshold of 0.08. PLS-SEM cannot replace the CORSIKA or GEANT4 simulations nor experimental measurements of the detector response. To this end, the first step for future research should be to gather expert responses and/or student responses using a validated questionnaire and follow-up measurement validation. The path coefficients, t values and p values, and confidence intervals should then be obtained using the SmartPLS bootstrapping [27], [28]. The structural model also ought to be re-designed in order to align directly with explicit paths and measurable endogenous outcomes (H1-H3). In addition it would be interesting to correlate the questionnaire constructs to the physical parameters: trigger efficiency, suppression factor, duty cycle, energy resolution, gamma-hadron separation and to collect the results of the questionnaire along with the CORSIKA and GEANT4 simulations, and, if possible, the TAIGA and NEVOD measurements. Finally, if actual responses to the real data are more complex, it might be worth considering whether R2 and R3 are more appropriately analysed as multidimensional constructs.

REFERENCES

- [1] M. Risse and P. Homola, "Search for ultra-high energy photons using air showers," *Mod. Phys. Lett. A*, vol. 22, no. 11, pp. 749–766, 2007, doi: 10.1142/S0217732307022864.
- [2] T. K. Gaisser, R. Engel, and E. Resconi, *Cosmic Rays and Particle Physics*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2016.
- [3] F. Schlüter and T. Huege, "Signal model and event reconstruction for the radio detection of inclined air showers," *J. Cosmol. Astropart. Phys.*, vol. 2023, no. 01, Art. no. 008, 2023, doi: 10.1088/1475-7516/2023/01/008.
- [4] H. J. Völk and K. Bernlöhr, "Imaging very high energy gamma-ray telescopes," *Exp. Astron.*, vol. 25, pp. 173–191, 2009, doi: 10.1007/s10686-009-9151-z.
- [5] Pierre Auger Collaboration, "Testing hadronic-model predictions of depth of maximum of air-shower profiles and ground-particle signals using hybrid data of the Pierre Auger Observatory," *Phys. Rev. D*, vol. 109, Art. no. 102001, 2024, doi: 10.1103/PhysRevD.109.102001.
- [6] R. D. Parsons and H. Schoorlemmer, "Systematic differences due to high energy hadronic interaction models in air shower simulations in the 100 GeV–100 TeV range," *Phys. Rev. D*, vol. 100,

Art. no. 023010, 2019, doi: 10.1103/PhysRevD.100.023010.

- [7] TAIGA Collaboration, "Optimization of electromagnetic and hadronic extensive air shower identification using the muon detectors of the TAIGA experiment," Nucl. Instrum. Methods Phys. Res. A, vol. 952, 2020, doi: 10.1016/j.nima.2018.12.045.
- [8] J. S. Pryga et al., "Analysis of the capability of detection of extensive air showers by simple scintillator detectors," Universe, vol. 8, no. 8, Art. no. 425, 2022, doi: 10.3390/universe8080425.
- [9] T. Huege, "Radio detection of cosmic ray air showers in the digital era," Phys. Rep., vol. 620, pp. 1–52, 2016, doi: 10.1016/j.physrep.2016.02.001.
- [10] F. G. Schröder, "Radio detection of cosmic-ray air showers and high-energy neutrinos," Prog. Part. Nucl. Phys., vol. 93, pp. 1–68, 2017, doi: 10.1016/j.pnpnp.2016.12.002.
- [11] S. Nonis, A. Leisos, and A. Tsigiotis, "Results and prospects of the Hellenic Open University air shower array," Phys. Sci. Forum, vol. 7, no. 1, Art. no. 7, 2023, doi: 10.3390/ECU2023-14015.
- [12] D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz, and T. Thouw, "CORSIKA: A Monte Carlo code to simulate extensive air showers," Forschungszentrum Karlsruhe, Karlsruhe, Germany, Rep. FZKA 6019, 1998.
- [13] S. Okukawa et al., "Neural networks for separation of cosmic gamma rays and hadronic cosmic rays in air shower observation with a large area surface detector array," Mach. Learn.: Sci. Technol., vol. 5, no. 2, Art. no. 025016, 2024, doi: 10.1088/2632-2153/ad3a33.
- [14] G. Mannocchi et al., "Multicomponent registration of the EAS," Bull. Russ. Acad. Sci.: Phys., vol. 83, no. 8, pp. 987–990, 2019, doi: 10.3103/S1062873819080033.
- [15] R. Ulrich, R. Engel, and M. Unger, "Hadronic multiparticle production at ultrahigh energies and extensive air showers," Phys. Rev. D, vol. 83, Art. no. 054026, 2011, doi: 10.1103/PhysRevD.83.054026.
- [16] R. Engel, D. Heck, and T. Pierog, "Extensive air showers and hadronic interactions at high energy," Annu. Rev. Nucl. Part. Sci., vol. 61, pp. 467–489, 2011, doi: 10.1146/annurev.nucl.012809.104544.
- [17] A. Pagliaro and A. La Barbera, "Machine learning for event reconstruction in imaging atmospheric Cherenkov telescopes," arXiv:2606.11741, 2026.
- [18] J. Albrecht et al., "The muon puzzle in cosmic-ray induced air showers and its connection to the Large Hadron Collider," Astrophys. Space Sci., vol. 367, Art. no. 27, 2022, doi: 10.1007/s10509-022-04054-5.
- [19] V. La Parola et al., "Machine learning-enhanced discrimination of gamma-ray and hadron events using temporal features: An ASTRI Mini-Array analysis," Appl. Sci., vol. 15, no. 7, Art. no. 3879, 2025, doi: 10.3390/app15073879.
- [20] Á. Pastor-Gutiérrez, H. Schoorlemmer, R. D. Parsons, and M. Schmelling, "Sub-TeV hadronic interaction model differences and their impact on air showers," Eur. Phys. J. C, vol. 81, no. 4, Art. no. 369, 2021, doi: 10.1140/epjc/s10052-021-09160-2.
- [21] I. Crawshaw, T. Yuan, E. Yildizci, L. Lu, and A. Fedynitch, "Probabilistic modeling of Cherenkov emission from particle showers," J. High Energy Phys., vol. 2026, no. 04, Art. no. 046, 2026, doi: 10.1007/JHEP04(2026)046.
- [22] H. Asorey and R. Mayo-García, "Calculation of the high-energy neutron flux for anticipating errors and recovery techniques in exascale supercomputer centres," J. Supercomput., vol. 79, pp. 8205–8235, 2023, doi: 10.1007/s11227-022-04981-8.
- [23] M. de Naurois and L. Rolland, "A high performance likelihood reconstruction of γ -rays for imaging atmospheric Cherenkov telescopes," Astropart. Phys., vol. 32, pp. 231–252, 2009, doi: 10.1016/j.astropartphys.2009.09.001.

- [24] R. K. Dey and A. Bhadra, "Search for gamma-ray induced showers from the lateral distribution of electrons in EAS," in Proc. 33rd Int. Cosmic Ray Conf. (ICRC), Rio de Janeiro, Brazil, 2013, arXiv:1503.06604.
- [25] K.-H. Kampert and M. Unger, "Measurements of the cosmic ray composition with air shower experiments," *Astropart. Phys.*, vol. 35, no. 10, pp. 660–678, 2012, doi: 10.1016/j.astropartphys.2012.02.004.
- [26] Original project research objectives and questionnaire supplied for the Hadron detector-response study.
- [27] J. F. Hair, G. T. M. Hult, C. M. Ringle, and M. Sarstedt, *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, 3rd ed. Thousand Oaks, CA, USA: Sage, 2022.
- [28] C. M. Ringle, S. Wende, and J.-M. Becker, *SmartPLS 4*. Bönningstedt, Germany: SmartPLS, 2024.
- [29] J. F. Hair, J. J. Risher, M. Sarstedt, and C. M. Ringle, "When to use and how to report the results of PLS-SEM," *Eur. Bus. Rev.*, vol. 31, no. 1, pp. 2–24, 2019, doi: 10.1108/EBR-11-2018-0203.
- [30] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," *Psychometrika*, vol. 16, no. 3, pp. 297–334, 1951, doi: 10.1007/BF02310555.
- [31] C. Fornell and D. F. Larcker, "Evaluating structural equation models with unobservable variables and measurement error," *J. Mark. Res.*, vol. 18, no. 1, pp. 39–50, 1981, doi: 10.2307/3151312.
- [32] J. Henseler, C. M. Ringle, and M. Sarstedt, "A new criterion for assessing discriminant validity in variance-based structural equation modeling," *J. Acad. Mark. Sci.*, vol. 43, no. 1, pp. 115–135, 2015, doi: 10.1007/s11747-014-0403-8.
- [33] L. Hu and P. M. Bentler, "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives," *Struct. Equ. Model.*, vol. 6, no. 1, pp. 1–55, 1999, doi: 10.1080/10705519909540118.
- [34] N. Kock and P. Hadaya, "Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods," *Inf. Syst. J.*, vol. 28, no. 1, pp. 227–261, 2018, doi: 10.1111/isj.12131.