Celebrating Computing in Astroinformatics: Exploring habitability of Exoplanets via Modeling and Machine Learning [2]

Snehanshu Saha, Ph.D

Center for Applied Mathematical Modeling, PES University, astrirg.org

Virtual Inst. of Astro-Particle Physics-Sequel to the talk by Dr. Margarita Safonova

19/10/2018

[2] Machine learning approach to understanding Astrophysical Data: The Final Frontier

Data Scientific Applications in Astronomy: Advent and Opportunities

- Modern astronomical instruments record huge volumes of data in the form of images, catalogs, raw data and signals in different bandwidths.
- A new wave of pursuits in astronomy thus involve the use of **statistics**, **machine learning**, and **artificial intelligence** to solve problems. An emerging **interdisciplinary** area of study which calls for scientists to collaborate from the fields of:
  1. Astronomy and astrophysics
  2. Statistics
  3. Mathematics
  4. Computer and Information sciences.

  Opportunities:
  1. to attract students from astrophysics, statistics, and computer science backgrounds; to create awareness in the fast developing field
  2. Scientometric analysis can provide insights to the pattern of collaboration between scientists from different domains, the popularity of various sub domains, and the direction in which future pursuits can

Our Work: Cobb-Douglas Habitability Production Function

The CD-HPF is a novel approach to estimate the habitability of an exoplanet. It is formulated as a minimization problem:

$$Y = f(R, D, T_s, V_e) = K \cdot R^{\alpha} \cdot D^{\beta} \cdot T_s^{\gamma} \cdot V_e^{\delta}$$

Key aspects of this model:

1. Inspired by econometric models
2. Important factors of habitability can contribute differently for different planets
3. convergence can be proved
4. can be solved by a computer in the **log-linear** form
5. new result powered by the principle of mathematical induction ensuring global optima under additional parameters such as orbital velocity, eccentricity etc.

Video link

Classification of Exoplanets

The rate at which exoplanets are being discovered is increasing. With the scheduled launch of the JWST in 2018, automated methods of classification need to be explored.

As a part of this endeavor, we plan:

1. Explore the efficacy of various classification algorithms
2. Propose the best methods of classification based on linear separability of data, etc.
3. Simulate the growth of data and tested the efficacy of ML algorithms on artificially augmented datasets
4. propose new methods to handle under-represented classes

One of the aspects of the PHL-EC catalog is that it assigns an eccentricity of 0 if the eccentricity of the orbit of a planet cannot be estimated. This prevents eccentricity from directly being used as a feature in the CD-HPF as it is in a **multiplicative** form.

To tackle this, the following methods are being tried:

1. Modeling the eccentricity of planets using perturbation theory
2. Modeling the eccentricity of planets using quadratic and logarithmic regression

Having determined the eccentricities reliably, the CD-HPF can be extended as:

$$Y = k \cdot R^{\alpha} \cdot D^{\beta} \cdot T_s^{\gamma} \cdot V_e^{\delta} \cdot S_f^{\zeta} \cdot (E + \epsilon)^{\tau}$$

thus incorporating other important factors for habitability.

Metallicity effects on habitability- Computational DoE Problem

Target variable, y, habitability; factorial analysis needed to determine importance of metals toward habitability; n^K designs where K is number of levels refers to designs with K factors where each factor (metal) has n levels; levels are continuous posing a huge design problem. Requires novel metaheuristic, Multi-stage Memetic Algorithm (MSMA) to solve the problem.

- **MSMA Operators:** Four operators are defined and used in two stages- used to accomplish. Each stage requires croosover and mutaion operators, Ω_1 and Ω_2 for stage 1 and, Ω_3 and Ω_4 for stage 2 respectively. We optimize over the support points in stage 1 and optimize over the distribution of the support points in stage 2. Hence the name, Multi stage memetic algorithm.

Remarks: Metallicity effects on habitability- Computational DoE Problem

- The proposed design is to be optimized in multiple stages, where in one stage we optimize over the design points and in the other stage, we optimize over the distribution of the support(design) points.
- Ω_1 is defined as the transposition(crossover) operator in breeding strategy while Ω_2 is the mutation operator, used to choose and create mutants.
- generate random population with K design runs, where design points are sampled randomly from the design space and all weights are assigned the same.
- Information Matrix is used to evaluate fitness of population.

Novelty

1. Knowledge creation in **Machine Learning** and **computational Design of Experiments** (DoE)
2. Novel applications in highly sporadic and imbalanced data arising in exoplanet studies, classification, star quasar classification, nova classification
3. "small-data treatment": The treatment for classification of exoplanets requires reasonable **under-sampling** and **artificial augmentation**. This is the opposite big-data pursuit and **requires novel interventions**!
4. New **theorems and proofs** solidifying the analytical and statistical foundation
5. Finding hidden correlations in the complex astronomical Big Data
6. Standardization of meta-data for better science

Novel contributions at a glance

- Novel *convex optimization* based model for *habitability score*, more *powerful* and *general* compared to the existing model, ESI
- propose improved habitability models under constraints and greater parameters by introducing *stochastic fronter analysis* to tackle the problem of curvature violation
- New *habitability catalog*
- Novel **computational method-MSMA** to determine relevance and importance of factors in metallicity of exoplanets

Publication Details

Title	Published/ Presented	Year	Impact factor
New Habitability Score via Data Analytic Modeling	Journal of Astronomy and Computing	2016	2.2
Book chapter on Machine Learning Approaches for Supernova Classification	Handbook of Research on Applied Cybernetics and Systems Science, IGI Global	2016	
Comparative study in classification methods and exploration of Habitability Catalog: PHL-EC	MNRAS	2017	4.952
Proxima Centauri b: Theoretical Validation of Potential Habitability via CD-HPF	J.Astronomy and Computing	to be 2018	2.2

Publication Details

Title	Published/ Presented	Year	Impact factor
ASTROMLSKIT: Open source statistical Toolkit	Neighborhood Astronomy Meeting	2015	
CD-HPF: New Habitability Score via Data Analytic Modeling	Space Science Symposium	2016	
Machine Learning Done Right: A case study in Quasar-Star Classification	AISC Springer and	to be	1.7
Classification of Novae	MNRAS	to be submitted	4.952

Introduction

- Astronomy has suddenly become an immensely data-rich field, with numerous digital sky surveys across a range of wavelengths, with many Terabytes of pixels and with billions of detected sources due to advancement of technology and new advanced technological telescopes and other such instruments.
- The nature of the data is very complex which needs to be analyzed and interpreted.
- The main focus here is to deal with the complexity of data like missing value, balancing of dataset or wrong data value.
- The entire work is broadly classified into : cosmology and habitability exploration.

Problem Statement

To design and develop a machine learning system which will analyze astronomical data and reduce the complexities in **parameter estimation** of various cosmological objects.

- [Karim Pichara et al, 2013] Used Bayesian networks for predicting missing data and Random Forest classifier to classify stars.
- [B. L. Lago et al, 2010] used two traditional methods chi-square and complete-likelihood approaches and compared and found that likelihood approach gives more restrictive constraints on MLCS2k2 light-curve fitter and SALT2 light-curve fitter.
- [Louis N.Irwin,2014] used ESI (Earth Similarity Index), as a parameter which rates the similarity of exoplanet to Earth based on the mass, size & temperature and PHI (Planetary Habitability Index) as another parameter and finally come up with a third parameter(based on above 2), BCI(Biological Complexity Index) to provide a more complete form to predict life outside the earth.

- [Z.E.Musielak,2014] had recently a new observation in which, based on the radio waves emissions ,they are predicting the exo-moons.
- [Debray,Raymond,2013] tried to train a classifier on light curves, which were represented as a time series of data and various attributes corresponding to this were stored in a binary table. An open source machine learning algorithm was used with a variety of classification methods. They got good accuracy results but it was not well tested for large data sets and left with much more scope for improvements.
- [Schulze-Makuch D. et.al. 2011] in their work showed a two- tiered approach to assess the habitability of exoplanets. As one parameter ESI was considered and the other parameter was PHI(Planetary Habitability Index). These 2 indices were properly designed and formulated by them in order to assess the habitability but requires more information of important planetary parameters.

Comparative Study in Classification Methods and Exploration of Habitability Catalog: PHL-EC

- The focus of this work is to explore the potential & efficacy of various machine learning algorithms like kNN, Decision Tree, Random Forest, Support Vector Machine , Naïve Bayes' and Linear Discriminant Analysis to automate the classification of newly discovered exoplanets.
- These algorithms are provided in an integrated manner to analyze the data from PHLs Exoplanet Catalog.

IEEE computer society

- Data used for this work has been taken from Planetary Habitability Laboratory- Exoplanet Catalog, (PHL-EC), University of Puerto Rico.
- Data set can be downloaded from, http://phl.upr.edu/projects/habitable-exoplanets-catalog/data/database
- PHL-EC data set has 68 features and were having 3416 confirmed exoplanets at the time of writing the paper. Currently the database has 3667 planets last updates on 21st May 2017.
- The 68 features have 13 categorical and 55 continuous features.
- The data set combines measured and modeled parameters of various sources. Hence, provides a good metric for visualization and statistical analysis.

Class labels in the dataset

After thorough analysis of the dataset, we have identified two class labels:

- Planetary Class classifies planets based on thermal zone (hot, warm or cold) and it's mass (Asteroidan, Mercurian, sub-terrain, Terrain, super-terrain, Neptunian & Jovian).
- Habitability Class classifies planets based on only temperature. It is described as follows -
  1. Hypopsychroplanet - very cold (below -50^0 C)
  2. Psychroplanet - cold (-50^0 C to 0^0 C)
  3. Mesoplanet - warm (0^0 C to 50^0 C)
  4. Thermoplanet - hot (50^0 C to 100^0 C)
  5. Hyperthermoplanet - very hot (100^0 C to 150^0 C)

- In the year 2015, there were 3 planets classes based on their thermal properties. They were - Mesoplanet, Psychroplanet and non-habitable planet classes. Later in 2016, 2 more classes were added. They are **Thermoplanet** and **Hypopsychroplanet**.
- We are concerned with the planets belong to Psychroplanet and Mesoplanet class as planets of these classes can sustain life.

There are few shortcomings in the catalog like missing data for few features, wrong data value or unbalanced data class.

- **Missing Data** were handled by inserting the class-wise mean for continuous valued attributes and the mode for categorical-valued attributes. Only about 1% of data in the dataset (after removing unimportant attributes for processing) is missing.

- The dataset was first scraped in June,2015. 664 planets with known surface temperature were considered, out of which 9 were a part of Mesoplanet class and 7 in Psychroplanet class. The remaining planets were in non habitable class.
- smaller datasets were constructed by considering all the planets in Mesoplanet, Psychroplanet class and 10 random planets from non habitable class.
- Classification and testing were performed on theses smaller dataset

- Later in the month of May 2016, dataset was updated with 3411 entries; Out of which 24 planets belong to Mesoplanet, 13 in Psychroplanet and remaining 3374 in Non habitable class. The non-habitable planet class dominates the other two classes. The result of classification performed on an unbalanced dataset is shown below.

Table: Accuracy Results for Each Algorithm Executed on Unbalanced PHL-EC Data Set

Algorithm	Accuracy (%)
Naïve Bayes	98.7
Decision Tree	98.61
LDA	93.23
k-NN	97.84
Random Forest	98.7
SVM	97.84

How did we handle unbalanced data?

All 13 entries from psychroplanet class were considered and 13 random and unique entries from other 2 classes. we frame a smaller dataset with 39 entries which is balanced. Each data set was divided in the ratio of 9:4(training : testing) and 500 iterations of training and testing were performed on each such data set. 500 such datasets were framed for analysis. Overall 2,50,000 iteration of training & testing were performed for each classifier. below is the result of classification -

Table: Accuracy results for each algorithm executed on artificially balanced PHL-EC data set

Algorithm	Accuracy(%)
Random Forest	96.311
Decision Tree	94.542
Naïve Bayes	92.583
LDA	77.396
k-NN	68.607

Other method: generation of duplicate data using repeated kNN-SVM

As the Mesoplanet and Psychroplanet classes are dominated by non-habitable class. To balance data in all 3 classes, we have generated duplicate data for the Mesoplanet and Psychroplanet class.

- The data imputation is done by assuming a Poisson distribution of features, followed by cross-validating the data repeatedly using K-NN and SVM classification.
- Repeated kNN-SVM algorithm is used to label the classes of the newly generated data.
- The accuracy of kNN is seen to be 99% and SVM is 100%.

CD-HPF: New Habitability Score via Data Analytic Modeling

- This work focuses on two important metrics , Earth Similarity Index (ESI) and Planetary Habitability Index (PHI) to assess the habitability of a planet as specified by the astrophysicists.
- In this work , we have successfully suggested a new way of obtaining the ESI value and then using this calculated new ESI (called here as Habitability Score), we implemented KNN algorithm to classify the planet in habitable class or non habitable class.
- The new way identified by us for this work is Cobb-Douglas Habitability Production Function (CD-HPF), which has its root in famous Cobb Douglas function used mainly in Economics.

Earth Similarity Index (ESI)

- Earth's Similarity Index(ESI)- considers Earth as the reference frame for habitability and measure physical similarity of any planetary body with the Earth. Its value lies between 0(no similarity) & 1(reference value). Generally, any planetary body with ESI over 0.8 is considered Earth like. It is given as-

Where ESI_x is the ESI value of a planet for x property and x_0 is the Earth's value for that property. Final ESI value is obtained by taking geometric means of individual values , where w is the weighting component . 4 components are mainly used for ESI calculation – radius , density, surface temperature and escape velocity.

Radius & density are used to calculate interior ESI (ESI_i) and surface temperature and escape velocity for surface ESI(ESI_s).

Then geometric mean of both is taken to generate final ESI.

Planetary Habitability Index (PHI)

- Some researchers defined another parameter based on the chemical composition of the planet called as PHI . It is specified as-

$$PHI = (S.E.C.L)^{1/4}$$

where S defines a substrate, E the available energy, C the appropriate chemistry and L the liquid medium.

- PHI in this form lacks some of the important properties that may be required to assess the habitability of a planet.

Proposed Approach

- A novel approach to analytically determine the habitability score of all confirmed exoplanets.
- Goal is to determine the likelihood of an exo-planet to be habitable using the newly defined habitability score (CDHS) based on Cobb-Douglas habitability production function (CD-HPF), which computes the habitability score by using measured and calculated planetary parameters.
- We looked for a feasible solution that maximizes habitability scores using CD-HPF with some defined constraints.
- CD-HPF is a feasible solution that maximizes the objective function, and is called an optimal solution under the constraints known as returns to scale.

CD-HPF: Optimization Function

- Returns to scale measure the extent of an additional output obtained when all input factors change proportionally. There are three types of returns to scale: 1. Constant Return to Scale (CRS) 2. Increasing Return to Scale (IRS) 3. Decreasing Return to Scale (DRS)

- CD-HPF uses the four parameters used in the ESI metric, i.e. surface temperature, escape velocity, radius and density to calculate the Cobb-Douglas Habitability Score (CDHS).
  - Analogous to the method used in ESI, two types of Cobb-Douglas Habitability Scores are calculated – the interior $CDHS_i$ and the surface $CDHS_s$. The final score is computed by a linear convex combination of these two, since it is well known that a convex combination of convex/concave function is also convex/concave.

IEEE computer society

- The interior $CDHS_i$, denoted by Y_1, is calculated using radius R and density D,

$$Y1 = CDHS_i = D^{\alpha}.R^{\beta}$$

- The Surface $CDHS_s$, denoted by Y_2, is calculated using surface temperature T_s and escape velocity V_e,

$$Y2 = CDHS_s = T_s^{\gamma}.V_e^{\delta}$$

- The final combination Y , which is the convex combination of Y_1 & Y_2 is given as-

$$Y = w' \cdot Y1 + w'' \cdot Y2\,,$$

where the sum of w' and w'' equals 1 and their values are weight to interior and surface CDHS. The final CDHPF function can be written as-

$$\mathbb{Y} = f(R, D, T_s, V_e) = (R)^{\alpha} \cdot (D)^{\beta} \cdot (T_s)^{\gamma} \cdot (V_e)^{\delta}\,,$$

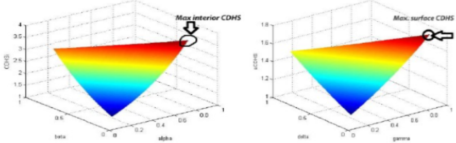


Figure: Interior and surface CDHS for DRS

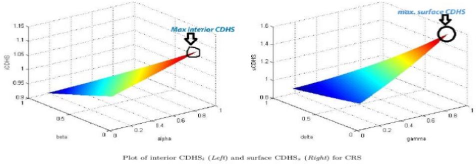


Figure: Interior and surface CDHS for CRS

- The computed CDHS is classified into 5 classes using Attribute-enhanced kNN algorithm.
- The probabilistic herding and thresholding are used to group the exo-planets according to their Y scores.
- Each CDHS value is compared with its K (specified by the user) nearest exoplanet's (closer Y values) CDHS value, and the class which contains maximum nearest to the new one is allotted as a class for it.

- Training data set is uniformly distributed between 5 classes, so that bias in the training set can be removed.
- Initially, each class holds one fifth of the training data and a new class, i.e. Class 6, defined as Earth's Class (or "Earth-League"), is derived by the proposed algorithm from first 5 classes where it contains data based on probabilistic herding and thresholding.

Normalization of CDHS

- CDHS value was also normalized for both CRS & DRS case to have the score in the range of 0-1 for all 664 exoplanets.
- We have applied Attribute-enhanced kNN algorithm to classify the planets. Under class 6 (Earth League Class), 16 exoplanets were categorized by the algorithm.

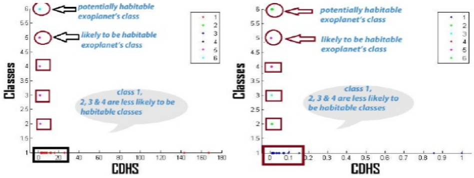


Figure: Results for Normalized & non-normalized cases of classification for DRS

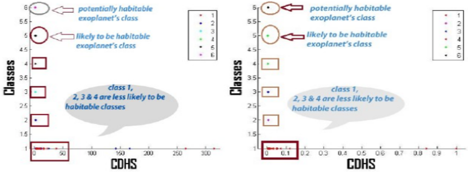


Figure: Results for Normalized & non-normalized cases of classification for CRS

Perturbed Cobb Douglas Model

- The idea is to include some more important parameters along with radius, density, surface temperature and escape velocity to the CDHS model.
- Two other important parameters i.e. stellar flux and eccentricity are added.
- For majority of the planets, eccentricity is assumed to be zero in the data set, which can't be used in Cobb Douglas formula (since it is a production function).
- Some changes are suggested -

$$\mathbb{Y} = f\left(R, D, T_s, V_e, S_f, E\right) = (R)^{\alpha} \cdot (D)^{\beta} \cdot (T_s)^{\gamma} \cdot (V_e)^{\delta} \cdot (S_f)^{\zeta} \cdot (E + \epsilon)^{\tau} \ .$$

where ϵ is the smallest eccentricity assumed in case eccentricity is zero (spherical orbit and hence stable) for that planet. We need to optimize eccentricity by using optimal control theory.

- We explore the efficacy of using a novel activation function in Artificial Neural Networks (ANN) in characterizing exoplanets
- We call this Saha-Bora Activation Function (SBAF) as the motivation - The function is demonstrated to possess nice analytical properties doesn't seem to suffer from local oscillation problems.
- Neural networks, commonly known as Artificial Neural network(ANN), is a system of interconnected units organized in layers, which processes information signals by responding dynamically to inputs. Layers of the network are oriented in such a way that inputs are fed at input layer and output layer receives output after being processed at neurons of one or more hidden layers.

- Hidden layers consist of computing neurons that are connected to input and output layers through a system of weighted connections. The network has ability to learn from input patterns, whereby with every input fed to the network, weights are updated in such a way that the error between the desired and observed output is minimum. Hidden layers are equipped with a special function called activation function to trigger neurons to process and propagate outputs across the network.
- A special class of ANN called Back propagation deals with computing the error between observed and desired output and later feeds this error back to the network with each cycle or 'epoch'. The weights are updated correspondingly and learning or training of the network is performed till the error is minimized–Activation function–functional mapping between inputs and outputs.

$$
\begin{aligned}
y &= \frac{1}{1+kx^{\alpha}(1-x)^{1-\alpha}};\\
\Rightarrow \ln y &= \ln 1 - \ln(1+kx^{\alpha}(1-x)^{1-\alpha})\\
&= -\ln(1+kx^{\alpha}(1-x)^{1-\alpha})\\
\Rightarrow \frac{1}{y}\frac{dy}{dx} &= -\frac{1}{(1+kx^{\alpha}(1-x)^{1-\alpha})}\cdot\Big[k\alpha x^{\alpha-1}(1-x)^{1-\alpha} - kx^{\alpha}(1-\alpha)(1-x
\\
&= -\frac{k}{(1+kx^{\alpha}(1-x)^{1-\alpha})}\cdot\Big[\alpha x^{\alpha-1}(1-x)^{1-\alpha} - (1-\alpha)x^{\alpha}(1-x)^{-\alpha}
\\
\Rightarrow \frac{dy}{dx} &= y\left[\frac{\alpha}{x} - (1-\alpha)\frac{1}{1-x}\right]kx^{\alpha}(1-x)^{1-\alpha}
\end{aligned} \tag{1}
$$

$$
\begin{aligned}
&= y\left[\frac{\alpha(1-x)-(1-\alpha)x}{x(1-x)}\right]kx^{\alpha}(1-x)^{1-\alpha} \\
&= y^2\left[\frac{\alpha-x}{x(1-x)}\right]kx^{\alpha}(1-x)^{1-\alpha}
\end{aligned} \tag{2}
$$

From the definition of the function, we have:

$$
\begin{aligned}
y &= \frac{1}{1+kx^{\alpha}(1-x)^{1-\alpha}} \\
\Rightarrow kx^{\alpha}(1-x)^{1-\alpha} &= \frac{1-y}{y}
\end{aligned} \tag{3}
$$

Substituting Equation 3 in 2, we obtain the final form

IEEE computer society

- $$\frac{dy}{dx} = y^2 \cdot \frac{\alpha - x}{x(1-x)} \cdot \frac{1-y}{y} = \frac{y(1-y)}{x(1-x)} \cdot (\alpha - x) \tag{4}$$

- Existence of Optima: Second order Differentiation of SBAF for Neural Network

$$\Rightarrow \frac{d^2y}{dx^2} = \frac{x(1-x) \cdot y(y-1)}{(x(1-x))^2} = \frac{y(y-1)}{x(1-x)}$$

The first derivative vanishes when $\alpha = x$, the second derivative is positive when $\alpha > x$ and is negative when $\alpha < x$

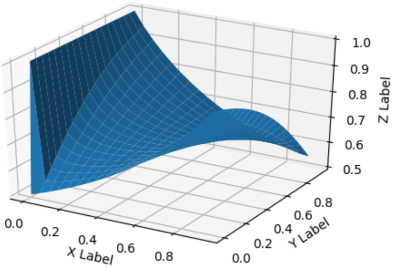


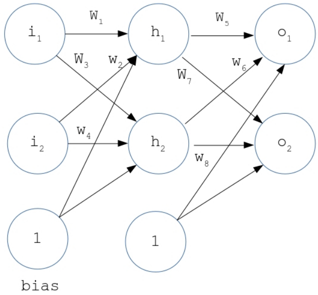
Figure: Surface Plot of SBAF

Remarks About SBAF

- This is equivalent to the CDHS formulation when CD-HPF is written as $y = kx^{\alpha}(1-x)^{\beta}$ where $\alpha + \beta = 1, 0 \leq \alpha \leq 1, 0 \leq \beta \leq 1$, k is suitably assumed to be 1 (CRS condition), and the representation ensures global maxima (maximum width of the separating hyperplanes) under such constraints. The new activation function to be used for training a neural network for habitability classification boasts of an optima.
- Evidently, from the graphical simulations, we observe less flattening of the function and therefore the formulation should be able to tackle local oscillations more easily as compared to the more generally used sigmoid function. Moreover, since $0 \leq \alpha \leq 1, 0 \leq x \leq 1, 0 \leq 1-x \leq 1$, the variable term in the denominator of SBAF, $kx^{\alpha}(1-x)^{1-\alpha}$ may be approximated to a first order polynomial. This may help us in circumventing expensive floating point operations without compromising the precision.

NN Architecture

The basic structure of the neural network consists of input layer, hidden layer and output layer. Let us assume the nodes at input layer are i_1, i_2 , at hidden layer h_1, h_2 and at output layer o_1, o_2.



Key Takeaways from SBAF

IEEE computer society

- Bingo! Perfect classification- 100% of the three classes
- Accuracy remains in tact even after removing surface temperature from the set of features
- Other activation functions are special cases of SBAF-namely **Sigmoid** and **ReLU**
- Maxima is unique in the defined interval. This will circumvent the local maxima problem.
- Production Function perspective- K can't be negative! This is useful for training and tuning the neural network.

Further Investigations in to the habitability metrics: PSO for Range of Scores

- Particle Swarm Optimization (PSO) is an optimization technique that iteratively improves a large number of randomly initialized solutions, called particles, to yield a globally best solution for a given problem
- It makes no assumptions of the problem being solved
- PSO does not require a gradient to be estimated and therefore does not require the objective function to be differentiable

Goal: Constant Elasticity Earth Similarity Approach (CEESA) score–maximize a Constant Elasticity of Substitution (CES) production function constructed with the density (D), radius (R), escape velocity (V_e), mean surface temperature (T_s) and the orbital eccentricity (E) of the planet under either Constant or Decreasing Returns to Scale to estimate similarity to Earth.

Problem Statement: maximizing the objective function,

$$Y = (r.R^\rho + d.D^\rho + t.{T_s}^\rho + v.{V_e}^\rho + e.E^\rho)^{\frac{\eta}{\rho}} , \quad (5)$$

where, $0 < \rho \leq 1$, coefficients r, d, t, v, e lie between 0 and 1 and add up to 1, and η is constrained by the scale of production used, $0 < \eta < 1$ under DRS and $\eta = 1$ under CRS.

Solution: Use the PSO algorithm to maximize Equation 5 to produce an Earth similarity metric.

Challenges:

- PSO was not designed to handle constraints in its classical definition.
- The algorithm needed to be modified to operate in a constrained search space

Solution Approach: CESSA-PSO

- PSO does not use the gradient of the objective function
- it must be able to simulate the gradient in order to gauge whether or not it is generating better solutions at the end of each iteration
- we can observe the value of the input variables as it pilots the objective to converge to a globally optimal solution.
- when particles are initialized or updated, the algorithm does not ensure the resulting solutions are feasible. The solution is twofold.
- resample each random solution from the uniform distribution until every initial solution is feasible
- while updating velocities always update toward a feasible solution, gathered so far by the algorithm, closest to the particle under update. This ensures that every particle eventually converges toward feasible solutions even if they do not necessarily traverse the feasible solution space.

Each iteration can be summed up as,

$$v_i = \omega . v_i + k_g (gbest - p_i) + k_p (lbest_i - p_i) \quad (6)$$
$$p_i = p_i + v_i \quad (7)$$

where ω is a constant in $(0, 1)$ and k_g, k_p are uniformly generated random numbers. These values function as inertial weights.

- Leaders in Optimization
- Simulate gradients
- Generating a score interval for planets–**BIG** difference from the earlier computational approaches!

Snapshot: Outcome of Models and ML Techniques

- Convergence of two approaches- Scoring via *Modeling* (CDHS/CESSA) with Feature based Classification via *Machine Learning* (SBAF)
- Habitability Score is not a rigid number- rather flexible!
- Existing habitability metrics (ESI/PHI) are specials cases of our metric CDHS
- Factoring eccentricity in to the habitability model

****

- Our Papers: http://astrirg.org/projects.html
- Our Codes: http://pesitsouthcompsoc.org/resources
- https://github.com/orgs/NeuralFuzzy/dashboard

T. M. Davis, E. Mortsell, J. Sollerman, A. C. Becker, S. Blondin, P. Challis, A. Clocchiatti, A. V. Filippenko, R. J. Foley, P. M. Garnavich, S. Jha, K. Krisciunas, R. P. Kirshner, B. Leibundgut, W. Li, T. Matheson, G. Miknaitis, G. Pignata, A. Rest, A. G. Riess, B. P. Schmidt, R. C. Smith, J. Spyromilio, C. W. Stubbs, N. B. Suntzeff, J. L. Tonry and W. M. Wood-Vasey 2007, Scrutinizing Exotic Cosmological Models Using ESSENCE Supernova Data Combined with Other Cosmological Probes. Astrophys.J, 666:716-725, DOI: 10.1086/519988

Riess et al. 2007, New Hubble Space Telescope Discoveries of Type Ia Supernovae at z ¿ 1: Narrowing Constraints on the Early Behavior of Dark Energy. Astrophys.J, 659:98-121, DOI: 10.1086/510378

Wood-Vassey et al. 2007, Observational Constraints on the Nature of the Dark Energy: First Cosmological Results from the ESSENCE Supernova Survey, Astrophys.J, 666:694-715, DOI: 10.1086/518642

Type Ia supernova data used by Davis, Mrtsell, Sollerman, et al. 2007. Retrieved from http://dark.dark-cosmology.dk/~tamarad/SN/

Thank You

Je vous remercie