

Hydrology मे Artificial Neural Network (Artificial Neural Networks in Hydrology)

क्षितिज मिश्रा¹, दीपक गुप्ता²

(Kshitij Mishra¹, Deepak Gupta²)

^{1,2}Computer Science & Engineering, NIT Arunachal Pradesh, Yupia 791112, India

Abstract

इस श्रृंखला में, Hydrology में Artificial Neural Network (ANN) की भूमिका की जांच करी गयी है। ANN लोकप्रियता प्राप्त कर रहे हैं, जैसा कि इस विषय पर विशेष रूप से पिछले दो दशकों में Hydrology पत्रिकाओं में दिखाई देने वाले कागजात की बढ़ती संख्या से प्रमाणित है। Hydrologic अनुप्रयोगों के संदर्भ में, यह Modelling उपकरण अभी भी अपने नवजात चरण में है। एक वैकल्पिक Modelling उपकरण के रूप में ANN की क्षमता के बारे में अभ्यास करने वाला Hydrologic समुदाय जागरूक हो रहा है। इस paper का उद्देश्य hydrologist के लिए ANN के परिचय के रूप में कार्य करना है। ANN के विभिन्न पहलुओं के विवरण और उनके उपयोग पर कुछ दिशानिर्देशों के अलावा, यह पत्र ANN की प्रकृति और Hydrology में अन्य Modelling दर्शनों की एक संक्षिप्त तुलना प्रदान करता है। ANN की ताकत और सीमाओं पर चर्चा और Physical model जैसे अन्य Modelling दृष्टिकोणों के साथ इसकी समानताएं पर भी चर्चा की गयी है।

Keywords: Artificial Neural Network (ANN); Hydrology; Back Propagation; Conjugate Gradient Descent; Radial Basis Function

1. पृष्ठभूमि

इस पेपर का विषय मूल प्रकृति का है और उन जलविज्ञानीओं के लिए लक्षित है जो अनिवार्य रूप से इस क्षेत्र में शुरुआती हैं। McCulloch and Pitts, 1943 द्वारा Artificial Neural Network का विकास (ANN) लगभग 70 साल पहले शुरू हुआ, जो मानव मस्तिष्क को समझने की इच्छा से प्रेरित थे और इसके कामकाज का अनुकरण करना चाहते थे। पिछले दशक के भीतर, अधिक परिष्कृत algorithm के विकास और शक्तिशाली गणना उपकरण के उद्भव के कारण ANN ने एक विशाल पुनरुत्थान का अनुभव किया है। Artificial Neural Network (ANN) computational tools एक बहुविकल्पीय input space से दूसरे में acquisition, representation और mapping की गणना करते हैं (Wasserman and P. D., 1989), इन tools की क्षमता की जांच के लिए व्यापक शोध समर्पित किया गया है। गणितीय रूप से, ANN अक्सर एक सार्वभौमिक अनुमानक के रूप में देखा जाता है। दिए गए pattern से रिश्ते की पहचान करने की क्षमता के कारण ANN के लिए pattern की पहचान, गैर-रेखिक जैसी जटिल Modelling समस्याएं, वर्गीकरण, association, और नियंत्रण बड़े पैमाने पर हल करना संभव हो जाता है। हालांकि 70 साल पहले, McCulloch, W. S., and Pitts (1943) द्वारा Artificial Neural Network का विचार प्रस्तावित किया गया था, पर iterative auto-associable neural network में Hopfield के प्रयास (Hopfield, 1982) के कारण ANN के विकास तकनीकों ने पिछले दशकों में पुनर्जागरण का अनुभव किया है। इस computational तंत्र की रूचि में जबरदस्त वृद्धि हुई, जब neural network के लिए Rumelhart et. al., 1986 ने Back Propagation algorithm विधि की फिर से खोज की। नतीजतन, ANN को विविध क्षेत्रों में अनुप्रयोग मिला है जैसे कि Neurophysiology, physics, Bio-medical engineering, Electrical engineering, Computer Science, acoustics, Cybernetics, Robotics, image processing, financing, and others. शुरुआती नब्बे के दशक से, Hydrology से संबंधित क्षेत्रों में ANN का सफलतापूर्वक उपयोग किया गया है जैसे Rain-runoff Modelling, stream flow forecasting, ground-water Modelling, water quality, water management policy, precipitation forecasting, hydrologic time series and reservoir operations. इस पेपर का लक्ष्य Artificial Neural Network का संक्षिप्त विवरण देना है, आमतौर पर इस्तेमाल किए गए algorithm का सारांश और Hydrologic समस्याओं के लिए ANN लागू करने के लिए दिशानिर्देश, ANN और अन्य Modelling दृष्टिकोण के बीच समानताएं और मतभेदों का वर्णन करना, और उनकी ताकत और सीमाओं पर चर्चा करना है।



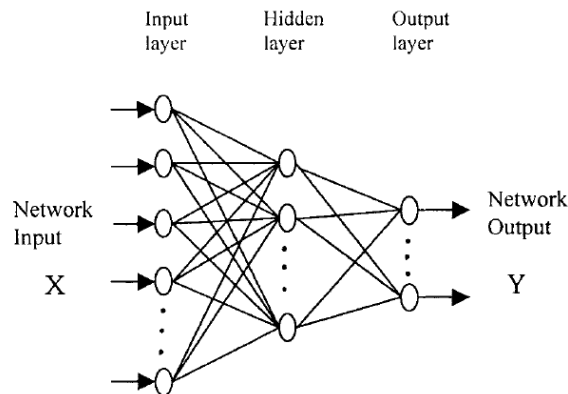
2. Artificial Neural Network का परिचय

ANN एक massively parallel-distributed information processing system है जिसमें प्रदर्शन विशेषताये मानव मस्तिष्क के जैविक neural network जैसी दिखती हैं। ANN को मानव संज्ञान या neural biology के गणितीय model के सामान्यीकरण के रूप में विकसित किया गया है। इसका विकास निम्नलिखित नियमों पर आधारित है:

- सूचना प्रसंस्करण कई एकल मूलत्वो पर होता है, जिन्हे Node, units, cells या neurons भी कहा जाता है।
- connection link के माध्यम से nodes के बीच संकेत पारित किए जाते हैं।
- प्रत्येक connection link में एक संबंधित weights या संख्यात्मक मानक होता है जो इसकी connection शक्ति दर्शाता है।
- प्रत्येक node output सिग्नल को निर्धारित करने के लिये network input पर एक गैर-रैखिक परिवर्तन लागू करता है, जिसे activation function कहा जाता है।

एक neural network इसके architecture द्वारा पहचाना जाता है जो nodes के बीच connection का pattern, connection weights निर्धारित करने की विधि, activation function दर्शाता है (Fausett and L., 1994)। एक सामान्य ANN में कई nodes होते हैं, जो एक विशेष व्यवस्था के अनुसार व्यवस्थित होते हैं। neural network को कई परतों में वर्गीकृत किया गया : Single (Hopfield network); bi-layer (Carpenter/Grossberg adaptive resonance network); और multi-layer (Back-propagation network)। ANN को सूचना प्रवाह और प्रसंस्करण की दिशा के आधार पर भी वर्गीकृत किया जा सकता है। feed forward network में, nodes आम तौर पर परतों में व्यवस्थित होते हैं, पहली input परत से शुरू और अंतिम output परत पर समाप्त होता है। कई छिपी परतें भी हो सकती हैं, और प्रत्येक परत में एक या अधिक nodes होते हैं। सूचना input से output पक्ष में गुजरती है। एक परत में nodes अगले में से जुड़े होते हैं, लेकिन उसी परत में नहीं। इस प्रकार, एक परत में node का output केवल पिछले परतों और संबंधित weights से प्राप्त input पर निर्भर होता है। यह पेपर feed forward network पर केंद्रित होगा, क्योंकि इन्हें आमतौर पर Hydrologic समस्याओं में उपयोग किया जाता है।

अधिकांश network में, input (पहली) परत उपलब्ध समस्या के लिए input variables (चर) प्राप्त करता है। इसमें सभी विशेषताएं होती हैं जो output को प्रभावित कर सकती हैं। input परत इस प्रकार पारदर्शी है और network की जानकारी प्रदान करने का माध्यम है। अंतिम या output परत में network द्वारा पूर्वानुमानित किए गए मान होते हैं और इस प्रकार model output को दर्शाते हैं। छिपी परतों की संख्या और प्रत्येक छिपी हुई परत में nodes की संख्या आमतौर पर trial-and-error प्रक्रिया द्वारा निर्धारित किया जाता है। network के पड़ोसी परतों के भीतर nodes पूरी तरह से link द्वारा जुड़े जाते हैं।



चित्र-1 Feed-forward तीन-परत ANN का configuration

Input-output relationship की भविष्यवाणी करने के लिए दोनों सिरों पर दो nodes की सापेक्ष connection ताकत को दर्शाने के लिए प्रत्येक link को एक synaptic weight सौंपा गया है। चित्र-1 feed forward network में तीन-परत ANN का configuration दिखाता है। इन प्रकार के ANN का उपयोग विभिन्न प्रकार की समस्याओं में किया जा सकता है, जैसे आंकड़ों को संग्रहित करना और याद करना, pattern वर्गीकृत करना, input pattern (space) से output pattern (space) तक सामान्य mapping करना, समान pattern को समूहीकृत करना, या constrained optimization problem के लिए समाधान ढूँढना। इस चित्र में, X एक system input vector है जो कई causal variables से बना है जो system व्यवहार को प्रभावित करता है, और Y system output vector है, जो system व्यवहार को दिखाने वाले resulting variables से बना है।

3. गणितीय पहलू

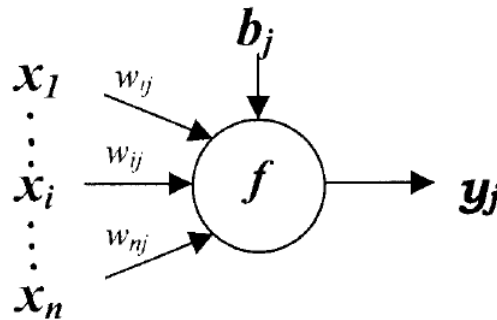
एक node का योजनाबद्ध आरेख चित्र नम्बर 2 में प्रदर्शित किया गया है। उस परत के आधार पर जिसमें node स्थित है, system input causal variables या अन्य nodes के output से आ सकता है। ये inputs एक input vector बनाते हैं $X = (x_1, \dots, x_i, \dots, x_n)$ । connection weights या मानको का अनुक्रम node के लिए weights vector $W_j = (w_{1j}, \dots, w_{ij}, \dots, w_{nj})$ देता है जहां w_{ij} पिछली परत के i^{th} node से इस परत के j^{th} node तक connection weight को दर्शाता है। node j का output y_j function f के मूल्य की गणना करके प्राप्त किया जाता है, जैसा कि नीचे दिया गया है:

$$y_j = f(X.W_j - b_j) \quad (1)$$

यहाँ पर b_j threshold value है, जिसे bias भी कहा जाता है, जो इस node से जुड़ा हुआ है। ANN व्याख्यान में, node के पूर्वाग्रह b_j को सक्रिय किए जाने से पहले बढ़ा लेना चाहिए। function f को activation function कहा जाता है। इसका functional रूप प्राप्त कुल input signal में node की प्रतिक्रिया निर्धारित करता है। सबसे अधिक इस्तेमाल किया जाने वाला $f(\cdot)$ sigmoid function है, जैसा कि नीचे दिया गया है:

$$f(t) = \frac{1}{1 + e^{-t}} \quad (2)$$

sigmoid function एक bounded, monotonic, non-decreasing function है जो एक वर्गीकृत, गैर-रैखिक प्रतिक्रिया प्रदान करता है। यह function किसी network को किसी भी गैर-रैखिक प्रक्रिया को मैप करने में सक्षम बनाता है। sigmoid function की लोकप्रियता आंशिक रूप से इसमें इस्तेमाल किये जाने वाले derivatives और इनकी simplicity के कारण है जिसका प्रयोग training प्रक्रिया के दौरान किया जाता है।



चित्र 2 : node j का योजनाबद्ध आरेख

4. Network Training

ANN को output vector $Y = (y_1, y_2, \dots, y_n)$ जो target vector $T = (t_1, t_2, \dots, t_p)$ के जितना संभव हो उतना करीब उत्पन्न करने के लिए एक training प्रक्रिया से गुजरना पड़ता है, जिसे learning भी कहा जाता है, और optimum weight vector matrix W और bias vector V पाये जाते हैं, जो आमतौर पर पूर्व निर्धारित error function को कम करता है। यह error function नीचे दर्शाया गया है:

$$E = \sum_p \sum_p (y_i - t_i)^2 \quad (3)$$

यहां, t_i वांछित output T का एक घटक है; y_i = संबंधित ANN output; p = output nodes की संख्या; और P = training pattern की संख्या। training एक ऐसी प्रक्रिया है जिसके द्वारा ANN के connection weights को निरंतर बढ़ाव प्रक्रिया के माध्यम से अनुकूलित किया जाता है। मुख्य रूप से दो प्रकार के training होते हैं - supervised और unsupervised। supervised training algorithm को training प्रक्रिया के मार्गदर्शन के लिए बाहरी शिक्षक की आवश्यकता होती है। यह आमतौर पर इंगित करता है कि training के लिए input और output के बड़ी संख्या में उदाहरण (या pattern) आवश्यक हैं। input system के causal variables होते हैं और output effect variables होते हैं। इस training प्रक्रिया में nodes के connection weights और threshold मानों का समायोजन और अनुकूलन पुनरावृत्ति ढंग से किया जाता है। training का प्राथमिक लक्ष्य connection function और threshold मानों

के set की खोज करके error function को कम करना है जो ANN के लिये लक्षित output के बराबर या करीब output उत्पन्न करने का कारण है। training पूरी होने के बाद, यह उम्मीद की जाती है कि दिए गए नए input के लिये ANN उचित परिणाम उत्पन्न करने में सक्षम है। इसके विपरीत, unsupervised training algorithm में शिक्षक शामिल नहीं होता है। training के दौरान, ANN को केवल एक input data set प्रदान किया जाता है, जो समान connection वाले वर्गों में उन input pattern को cluster करने के लिए अपने connection weights को स्वचालित रूप से अनुकूलित करता है। ऐसे अवसर होते हैं जब इन दो training रणनीतियों का संयोजन reinforcement training के रूप देखा जाता है। training pattern की श्रृंखला पर network प्रदर्शन को rate करने के लिए score या grade का उपयोग किया जाता है। अधिकांश जलविद्युत अनुप्रयोगों में supervised training का उपयोग किया गया है। ANN के nodes किस तरह संरचित हैं, यह प्रशिक्षित करने के लिए इस्तेमाल किए गए algorithm पर निर्भर करता है। इस शेष भाग में कई सामान्य रूप से उपयोग किए जाने वाले training algorithm शामिल हैं।

5. Back Propagation

Back Propagation, ANN training के लिए सबसे लोकप्रिय algorithm है [विस्तृत विवरण के लिए Wasserman and P. D., 1989 और Fauset and L., 1994 देखें]। यह अनिवार्य रूप से एक gradient descent तकनीक है जो network error function (3) को कम करता है। training data set का प्रत्येक input pattern input परत से output परत तक network के माध्यम से पारित किया जाता है। network output की वांछित target output के साथ तुलना की जाती है, और एक error की गणना (3) उसके आधार पर की जाती है। यह error network के माध्यम से प्रत्येक node में पीछे कि तरफ propagate किया जाता है, और संगत रूप से connection weights समीकरण के आधार पर समायोजित किया जाता है।

$$\Delta w_{ij}(n) = -\varepsilon * \frac{\partial E}{\partial w_{ij}} + \alpha \Delta w_{ij}(n-1) \quad (4)$$

यहां $\Delta w_{ij}(n)$ और $\Delta w_{ij}(n-1) = n^{th}$ और $(n-1)^{th}$ वें pass में node i और j के बीच में weights वृद्धि दर्शाते हैं। Bias मूल्यों के सुधार के लिए एक समान समीकरण लिखा गया है। (4) में, ε और α को क्रमशः learning rate और momentum कहा जाता है। momentum factor error सतह के बहुत ही सपाट क्षेत्रों में training तेज कर सकता है और weights में oscillations को रोकने में मदद करता है। Global minimum के बजाए local minimum में फंसे हुए training प्रक्रिया से बचने के लिए learning rate का उपयोग किया जाता है। Back Propagation algorithm में दो चरण शामिल हैं। पहला चरण एक forward pass है, जिसमें output परत तक पहुंचने के लिए input के प्रभाव को network के माध्यम से आगे बढ़ाया जाता है। error की गणना के बाद, network के माध्यम से दूसरा चरण backward pass शुरू होता है। output परत पर errors को input परत की तरफ संशोधित weights के साथ ((4) के अनुसार) वापस प्रसारित किया जाता है। Back Propagation steepest gradient descent के आधार पर first order विधि है, जिसमें दिशा vector gradient vector के नकारात्मक के बराबर set किया जाता है। नतीजतन, न्यूनतम error स्थिति तक पहुंचने का प्रयास करते समय समाधान अक्सर zig-zag path का पालन करता है, जो training प्रक्रिया को धीमा कर सकता है। learning rate के उपयोग के बावजूद local minimum में training प्रक्रिया का फंसना संभव है। अगले खंड में, Conjugate Gradient Descent का वर्णन किया गया है जो इस समस्या को कम करने में मदद कर सकता है।

6. Conjugate Gradient Descent

Conjugate Gradient Descent पहली बार Fletcher and Reeves (1964) द्वारा सामान्य अनियंत्रित अनुकूलन समस्याओं पर प्रयोग की गयी। Back Propagation के विपरीत, Conjugate Gradient Descent error gradient की दिशा के साथ आगे नहीं बढ़ती है, लेकिन पिछले pass की दिशा के orthogonal दिशा में बढ़ती है। यह भविष्य के चरणों को वर्तमान चरण के दौरान प्राप्त न्यूनतमकरण को प्रभावित करने से रोकता है।

Fletcher and Reeves (1964) ने दिखाया है कि कोई भी न्यूनतमकरण विधि Conjugate Gradient Descent द्वारा विकसित की गयी quadratically convergent है। यदि यह गैर-वर्गिक पुनरावृत्ति समस्या के मामले में उपयोग किया जाता है, जैसे error समीकरण (3), एक convergent मानदंड की आवश्यकता होती है। यदि $P(n)$ Back Propagation के n^{th} पुनरावृत्ति पर दिशा vector को इंगित करने के लिए प्रयोग किया जाता है, (4) को फिर से लिखा जा सकता है:

$$W(n+1) = W(n) + \varepsilon P(n) \quad (5)$$

यहां $W(n+1)$ और $W(n) = (n+1)^{th}$ और n^{th} पुनरावृत्ति में कुछ node के weight vector, $\varepsilon =$ learning rate। प्रारंभिक दिशा vector नकारात्मक gradient vector $g(n)$ प्रारंभिक बिंदु $n=0$ पर, यानी $P(0) = -g(0)$ के

बराबर set है। प्रत्येक क्रमिक दिशा vector की वर्तमान gradient vector और पिछले दिशा vector के एक रैखिक संयोजन द्वारा गणना की जाती है जैसे कि नीचे दिया गया है:

$$P(n+1) = -g(n+1) + \beta(n)P(n) \quad (6)$$

यहां, $\beta(n)$ एक समय-निर्भर parameter है, जो Fletcher and Reeves (1964) ने नीचे गये सूत्र के अनुसार परिभाषित है:

$$\beta(n) = \frac{g^T(n+1)g(n+1)}{g^T(n)g(n)} \quad (7)$$

Conjugate Gradient Descent algorithm कई मामलों में training प्रक्रिया को तेज कर सकता है।

7. Radial Basis Function

एक Radial Basis Function (RBF) network को तीन-परत network के रूप में माना जा सकता है जिसमें छिपी परत बिना किसी समायोज्य parameter के एक निश्चित गैर-रैखिक परिवर्तन करता है (Leonard et al., 1992)। इस परत में कई nodes और parameter vector होते हैं जिन्हें “केंद्र” कहा जाता है, जिसे छिपी हुई परत के weight vectors माना जा सकता है। मानक euclidean दूरी का उपयोग यह मापने के लिए किया जाता है कि केंद्र से input vector कितना दूर है। प्रत्येक node के लिए, network input के input vector और केंद्र के बीच euclidean दूरी की गणना एक गैर-रैखिक function द्वारा की जाती है जो छिपी हुई परत में nodes के output को निर्धारित करती है। output परत तब इन परिणामों को एक रैखिक रूप में जोड़ती है। एक RBF network के output y की नीचे दिये गये समीकरण द्वारा गणना की जा सकती है:

$$y = f(u) = \sum_{i=1}^n w_i R_i(x) + w_0 \quad (8)$$

यहां w_i छिपे हुए neuron और output neuron के बीच connection weights; w_0 = bias; x = input vector। functions $R_i = R^n \Rightarrow R$ Radial Basis Function हैं जिनका सामान्य रूप है:

$$R_i(x) = \phi(\|x - c_i\|) \quad (9)$$

$\phi(\cdot)$ का मूल्य उद्गम पर अधिकतम होता है और तेजी से क्षय होता है जैसे जैसे इसका argument infinity की ओर जाता है। जैसे जैसे euclidean दूरी input vector और केंद्र के बीच में बढ़ती है $\phi(\cdot)$ function को शून्य तक पहुंचने की आवश्यकता होती है। $\phi(\cdot)$ के लिए विभिन्न विकल्प उपलब्ध हैं, पर सामान्यतः Radial Basis Function gaussian function द्वारा वर्णित है।

$$R_i = -\exp\left(-\sum_{j=1}^n \frac{\|x_j - c_{ij}\|^2}{2\sigma_{ij}^2}\right) \quad (10)$$

यहाँ $c_i^T = [c_{i1}, c_{i2}, \dots, c_{im}]$ ग्रहणशील क्षेत्र का केंद्र है; और $\sigma_{ij} =$ gaussian function की चौड़ाई। RBF network रचना का प्रमुख कार्य, केंद्र c को निर्धारित करना है। training set से केंद्रों को यादृच्छिक रूप से चुनने का सबसे आसान तरीका हो सकता है। दूसरा दृष्टिकोण input training set को कई समूह में k-means तकनीक का उपयोग करके clustering करना है और प्रत्येक समूह के केंद्र को केंद्र के रूप में चुनना है। इसके अलावा, c को w_i के साथ network parameter के रूप में माना जा सकता है और error-सुधार training के माध्यम से समायोजित किया जा सकता है। केंद्र निर्धारित होने के बाद, छुपी परत और output परत के बीच connection weight w_i सामान्य Back Propagation training के माध्यम से निर्धारित किया जा सकता है। RBF network और Back-propagation के बीच प्राथमिक अंतर छुपे nodes से जुड़े गैर-रैखिकताओं की प्रकृति में है। Back Propagation में गैर-रैखिकता एक निश्चित function जैसे sigmoid द्वारा कार्यान्वित किया जाता है। दूसरी तरफ RBF विधि training set में data पर अपनी गैर-रैखिकता का आधार बनाती है। एक बार छुपी परत में सभी आधार functions को पाने के बाद, output परत पर network को केवल रैखिक योग विधि में सीखने की जरूरत है।

8. Hydrology में ANN और अन्य Modelling विधियाँ

जल विज्ञान (Hydrology) पृथ्वी की सतह, मिट्टी और वातावरण पर पानी और इसकी गुणधर्म, वितरण और प्रभाव का वैज्ञानिक अध्ययन है। अधिकांश Hydrologic प्रक्रियाएं अत्यधिक गैर-रैखिक होती हैं और स्थानिक और लौकिक परिवर्तनशीलता की उच्च डिग्री प्रदर्शित करती हैं। वे parameter अनुमानों में अनिश्चितता से और जटिल हैं। जलविज्ञानी अक्सर पूर्वानुमान की समस्याएं जैसे कि run-off, वर्षा, प्रदूषक सांद्रता, और पानी के चरणों के आकलन की समस्याओं के साथ सामना करते हैं। इस प्रकार की जानकारी Hydrologic और hydrologic engineering design के साथ ही जल संसाधन प्रबंधन में आवश्यक है।

वर्तमान में जल विज्ञान में उपयोग किए जाने वाले model को व्यापक रूप से तीन श्रेणियों के तहत समूहीकृत किया जा सकता है: empirical, geomorphology based, और physically based. Empirical model black-box के रूप में Hydrologic system (जैसे watershed) को मानते हैं और ऐतिहासिक input (वर्षा, तापमान, आदि) और output (watershed run-off को stream gauge पर मापा जाता है) के बीच संबंध खोजने का प्रयास करते हैं। Lumped catchment model इस श्रेणी के अंतर्गत आते हैं। इन विधियों को लंबे ऐतिहासिक record की आवश्यकता होती है और उनके पास कोई भौतिक आधार नहीं होता है इसलिये ungauge watershed के लिए लागू नहीं होते हैं (ASCE Task Committee on Application of ANN in Hydrology, 2000)। इन प्रकार के model पर एक सुधार geomorphology based model है। ये model watershed संरचना और stream network को दर्शाते हैं, लेकिन अलग-अलग watershed इकाइयों (streams and overland sections) की प्रतिक्रिया की रैखिकता से संबंधित विभिन्न मानको को लेने की आवश्यकता है।

Physical based model 'भौतिकी (physics)' को दर्शाते हैं क्योंकि उस तरह से उन्हें सबसे अच्छी तरह समझा जाता है। आम तौर पर, उनमें partial differentiation समीकरणों की एक प्रणाली का समाधान शामिल होता है जो watershed के भीतर प्रवाह प्रक्रियाओं की हमारी सर्वोत्तम समझ को दर्शाते हैं। ज्यादातर समस्याओं के लिए, समय-स्थान आयामों को nodes के एक अलग set में विघटित करके एक संख्यात्मक समाधान की खोज की जाती है। इसका तात्पर्य यह है कि ऐसे model सर्वश्रेष्ठ काम करते हैं जब watershed की भौतिक विशेषताओं के data, model grid पैमाने पर उपलब्ध होते हैं। इस प्रकार का data भारी मात्रा में watershed शोध में भी शायद ही उपलब्ध है। स्थानिक परिवर्तनशीलता और watershed को सटीक तरीके से प्रस्तुत करने की समस्या और इनकी अनिश्चितता के कारण stochastic watershed model का प्रस्ताव दिया गया है। यह दिखाने के लिए कि watershed इकाइयों की प्रतिक्रिया काफी गैर-रैखिक है, ऐसे model का प्रदर्शन hypothetical अध्ययनों में किया गया है। ये model parameter अनुमान की पहचान, अनुमान और विशिष्टता जैसी समस्याओं से पीड़ित हैं। वर्तमान computing क्षमताओं के साथ भी, watershed में physical model द्वारा दिये गये मान बस एक अच्छे अनुमान है। भौतिक सिद्धांतों पर आधारित model ungauge watershed पर लागू होते हैं, लेकिन लगभग हमेशा, उन्हें model अंशांकन उद्देश्यों के लिए ऐतिहासिक data की आवश्यकता होती है। इन सीमाओं के बावजूद, उचित रूप से उपयोग किए जाने पर ये model कई Hydrology समस्याओं के लिए बहुत उपयोगी साबित हुए हैं। Engineers और hydrologists physically based model का सफलतापूर्वक उपयोग करके watershed system के जटिल व्यवहार को model करते हैं और मानको की भविष्यवाणी करते आये हैं। लेकिन यह स्पष्ट है कि यह समस्या अभी भी असुलभी है। शायद इस कारण से वैकल्पिक Modelling दृष्टिकोण अभी भी मांगा जा रहा है।

इस संक्षिप्त चर्चा के आधार पर, ANNs को empirical model के रूप में वर्गीकृत करना होगा। हम इस दृष्टिकोण को 'model' कहते हैं क्योंकि इसमें जलविज्ञान (Hydrology) में अन्य Modelling दृष्टिकोणों के साथ कई विशेषताएँ समान हैं। model चयन की प्रक्रिया उपयुक्त network architecture के निर्धारण के बराबर माना जा सकता है। इसी तरह, model training, सत्यापन training और परीक्षण के साथ model अंशांकन और सत्यापन की पहचान की जा सकती है। कई मामलों में, ANN Hydrology में regression आधारित model के समान होते हैं, सिवाय इसके कि उन्हें गणितीय रूप के विनिर्देश की आवश्यकता नहीं होती है। इसके अलावा, ANN छिपी परतों की संख्या और इनमें से प्रत्येक परत से जुड़े nodes के विकल्प के साथ उपलब्ध स्वतंत्रता के कारण अधिक बहुमुखी हैं। अतः ANN संरचना एक साथ कई पथों के साथ संसाधित होने की अनुमति देती है, जिससे समानांतर कार्यान्वयन के अवसर मिलते हैं।



9. कुछ अन्य समस्याएं

कुछ ANN अनुप्रयोगों ने उन्हें network पर पेश करने से पहले input / output मात्रा को scale करने के महत्व पर बल दिया है। उच्च गैर-रैखिकता प्रदर्शित करने वाली समस्याओं के लिए, variables 0 से 1, या कुछ अन्य उपयुक्त सीमा के बीच में आने के लिए scale किए जाते हैं। इस प्रकार की scaling समाधान स्थान को सुचारू बनाने और कुछ शोर प्रभावों को औसत करने के लिए होती है। हालांकि, इस प्रक्रिया के माध्यम से जानकारी खोने का कुछ खतरा है। weights और threshold मूल्यों को initialize करना एक महत्वपूर्ण समाधान है। जितना प्रारंभिक अनुमान optimum weight space के करीब है, उतनी training प्रक्रिया तेज है। हालांकि, अच्छे weights का प्रारंभिक अनुमान बनाने का कोई तरीका नहीं है, और वे एक यादृच्छिक fashion में शुरू किए जाते हैं। आमतौर पर, छोटे यादृच्छिक भार सुझाए जाते हैं।

10. ताकत और सीमाएं

ANN एक आकर्षक computational tool बनने के कुछ कारण निम्नलिखित हैं:

- बिना किसी भौतिक मानको के वे input और output variables के बीच संबंध पहचानने में सक्षम हैं।
- जब training set में noise और मापन errors होते हैं तब भी वे अच्छी तरह से काम करते हैं।
- वे बदलती परिस्थितियों की भरपाई के लिए समय के साथ समाधान के अनुकूल होने में सक्षम हैं।
- उनके पास अन्य अंतर्निहित सूचना-प्रसंस्करण विशेषताओं का अधिकार होता है और एक बार प्रशिक्षित करके इन्हे इस्तेमाल करना आसान होता है।

अक्सर, जल विज्ञान में, भौतिक रूप से आधारित तरीकों का उपयोग करके समस्याओं को स्पष्ट रूप से समझा नहीं जा सकता है या अर्थपूर्ण विश्लेषण के लिए ये गलत परिभाषित होते हैं। यहां तक कि जब ऐसे model उपलब्ध होते हैं, तब भी उन्हें कुछ मानको पर भरोसा करना पड़ता है; जो ANN को अधिक आकर्षक बनाता है। इसके अलावा, ANN नियमित रूप से जटिल partial differential समीकरणों को हल किए बिना अंतर्निहित प्रक्रिया की गैर-रैखिकता का model करते हैं। regression-आधारित तकनीकों के विपरीत, input और output के बीच संबंधों के गणितीय मूल्यों के मानको की जानकारी की आवश्यकता नहीं है। input और output में noise की उपस्थिति network के भीतर वितरित प्रसंस्करण की वजह से सटीकता के गंभीर नुकसान के बिना ANN द्वारा संभाली जाती है। यह, activation function की गैर-रैखिक प्रकृति के साथ, वास्तव में ANN की सामान्यीकृत क्षमताओं को बढ़ाता है और उन्हें जल विज्ञान में समस्याओं की एक बड़ी श्रेणी के लिए वांछनीय बनाता है। हालांकि कई अध्ययनों से संकेत मिलता है कि ANN जल विज्ञान में संभावित रूप से उपयोगी उपकरण साबित हुए हैं पर उनके नुकसान को नजरअंदाज नहीं किया जाना चाहिए। ANN अनुप्रयोगों की सफलता data की मात्रा और गुणवत्ता दोनों पर निर्भर करती है। इस आवश्यकता को आसानी से पूरा नहीं किया जा सकता है, क्योंकि कई Hydrologic record काफी अतीत तक नहीं जाते हैं। अक्सर, आवश्यक data उपलब्ध नहीं होता है और इसे और तरह के परीक्षण model जैसे अन्य माध्यमों से उत्पन्न किया जाना चाहिए। यहां तक कि जब लंबे ऐतिहासिक record उपलब्ध होते हैं, यह निश्चित नहीं होता कि इस समय अवधि में स्थितियां एकरूप रहेंगीं। इसलिए, मानव प्रणाली द्वारा अपेक्षाकृत स्थिर और अप्रभावित प्रणाली पर दर्ज data set वांछनीय हैं। अक्सर अस्थायी विविधता वर्तमान input के रूप में पिछले input / output द्वारा प्राप्त की जाती है। हालांकि, यह स्पष्ट नहीं है कि अस्थायी प्रभावों को शामिल करने के लिए अतीत में कितनी दूर जाना चाहिए। इससे परिणामस्वरूप ANN संरचना अधिक जटिल हो जाती है। ANN की एक और बड़ी सीमा भौतिक अवधारणाओं और संबंधों की कमी में है। यह इस पद्धति के प्रति संदिग्ध दृष्टिकोण प्राथमिक कारणों में से एक रहा है। यह तथ्य कि network architecture का चयन करने का कोई मानकीकृत तरीका नहीं है, भी आलोचना प्राप्त करता है। network architecture, training algorithm, और error की परिभाषा की पसंद आमतौर पर समस्या के भौतिक पहलुओं के बजाय उपयोगकर्ता के पिछले अनुभव और वरीयता द्वारा निर्धारित की जाती है।

11. निष्कर्ष

यह पेपर Artificial Neural Networks (ANNs) के परिचय के रूप में कार्य करता है जिसमें Hydrologic समस्याओं के लिए उनके अनुप्रयोग पर जोर दिया गया है। यह ANN का एक संक्षिप्त विवरण प्रस्तुत करता है, उसकी अंतर्निहित अवधारणा और गणितीय पहलू, और अन्य Modelling दृष्टिकोण के सापेक्ष में ANN की भूमिका को बताता है। कुछ लोकप्रिय ANN architecture और algorithm पर चर्चा की गयी है। Hydrology में समस्याओं के लिए ANN के अनुप्रयोग के लिए दिशानिर्देश प्रस्तुत किए गये हैं। इस पद्धति की योग्यता और कमियों पर चर्चा की गयी है। पेपर में दी गयी चर्चा के बाद यह निष्कर्ष दिया जा सकता है कि Hydrology में ANN model पारंपरिक रूप से इस्तेमाल किए जाने वाले physical model से बेहतर होते हैं और जटिल गैर-रैखिक data के साथ भी बेहतर परिणाम दे सकते हैं।

संदर्भ (References)

1. McCulloch, W. S., and Pitts, W. (1943). "A logic calculus of the ideas immanent in nervous activity."



- Bull. of Math. Biophys., 5, 115–133. MIT Press, Cambridge, Mass., 318–362.
2. Wasserman, P. D. (1989). *Neural computing: theory and practice*. Van Nostrand Reinhold, New York.
 3. Hopfield, J. J. (1982). “Neural networks and physical systems with emergent collective computational abilities.” *Proc., Nat. Academy of Scientists*, 79, 2554–2558.
 4. Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). “Learning internal representations by error propagation.” *Parallel distributed processing*, Vol. 1.
 5. Fausett, L. (1994). *Fundamentals of neural networks*. Prentice Hall, Englewood Cliffs, N.J.
 6. Fletcher, R., and Reeves, C. M. (1964). “Function minimization by conjugate gradient.” *Comp. J.*, 7, 149–154.
 7. Leonard J. A., Kramer, M. A., and Ungar, L. H. (1992). “Using radial basis functions to approximate a function and its error bounds.” *IEEE Trans. on Neural Networks*, 3(4), 624–627.
 8. ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. (2000). *Artificial neural networks in hydrology. I: Preliminary concepts*. *Journal of Hydrologic Engineering*, 5(2), 115-123.

