# Crops Yield – Climate Regression Modeling Based on AIC

Dr. Sarfaraz
Chief Met. PMD-Karachi
sarfarazmet@hotmail.com

# Sequence

- Pakistan main crops spatial mapping
- Akaike Information Criteria (AIC) Methodology
- AIC-based regression model for wheat-climate association
- AIC-based regression model for rice-climate association
- AIC-based regression model for cotton-climate association
- AIC-based regression model for maize-climate association
- Significant climate parameters identified for each crop
- Future drought scenario under animated global warming
- Potential impact on crops due to future projected climate
- Adaptation
- Conclusion

# Methodology for AIC (Akaike Information Criteria) Calculation

- The *AIC* estimates the comparative value of arithmetical prototypes for a particular dataset.
- The Akaike's procedures in essence are information-theoretic as they rely on the K-L information (Akaike, 1973, 1983b, 1992 & 1994).
- Having a set of candidate simulations, *AIC* is then computed for each simulation and the one with a minimal *AIC* score is regarded as the finest prototype for given empirical data.
- Model selection based on *AIC* is comparable to definite cross-validation procedures (Stone, 1974 & 1977).
- The *AIC* value for each model is calculated by

$$AIC = n*\ln(RSS/n)+2*K$$

where, 'n' is the total quantity of observation (sample size), K is the degree of freedom or number of parameters, and *RSS* is the *residual sum of squares*.

- Then using refinement technique for corrected estimate for small data samples (Hurvich & Tsai, 1989; Burnham & Anderson, 2002) the AIC-corrected, AICc, is calculated as;

$$AICc = AIC+(2*K(K+1))/ n-K-1$$

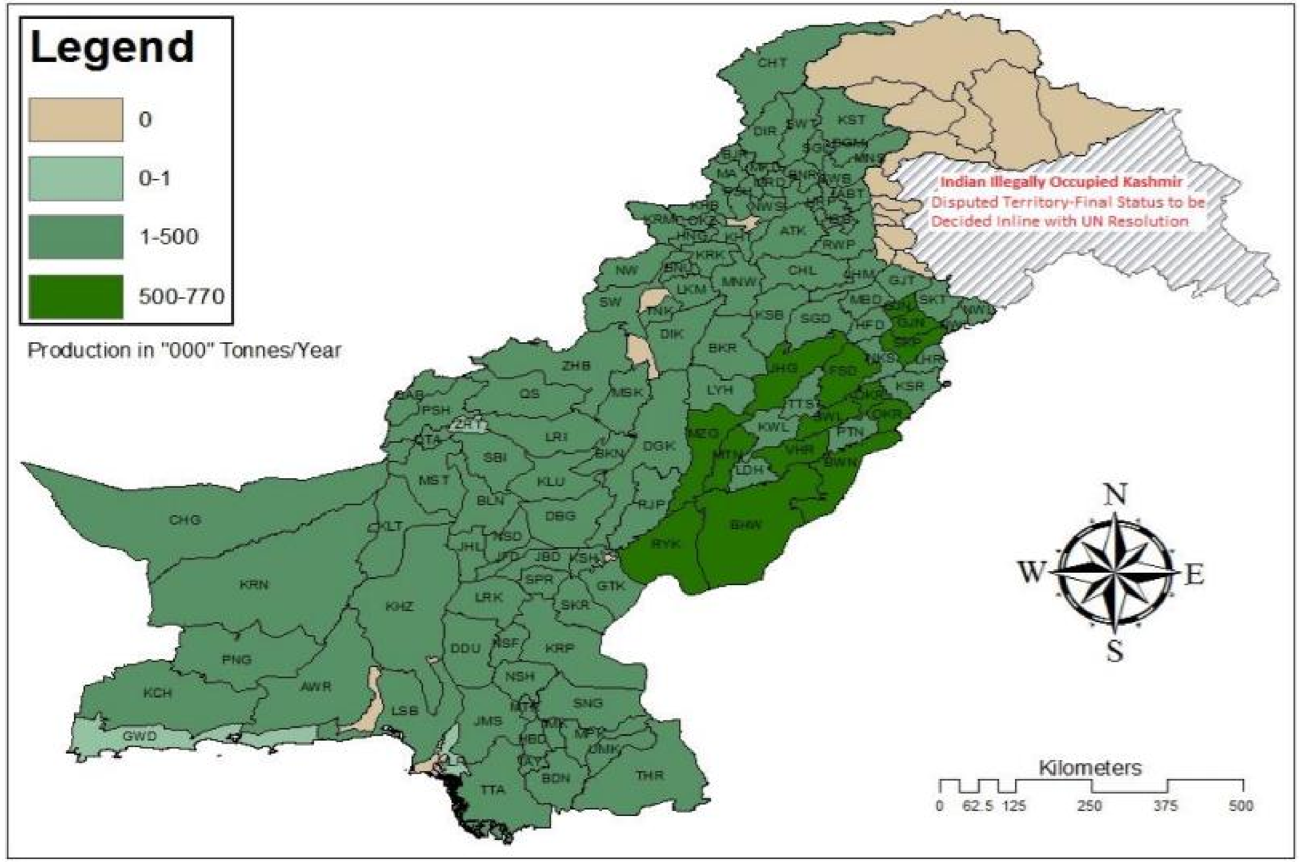
- And finally the shortest distance to the 'truth' Δi for each model is calculated by

$$\Delta i = AICi - minAICc$$

- Δi is the strength of evidence whose minimum value gives the best model (Burnham & Anderson, 2001).
- The crops-climate association or correlation is assessed by linear and multiple regression models using the statistical package for social scientists, SPSS (IBM SPSS Statistics 21, 2012).

# District-wise *Wheat* production across Pakistan

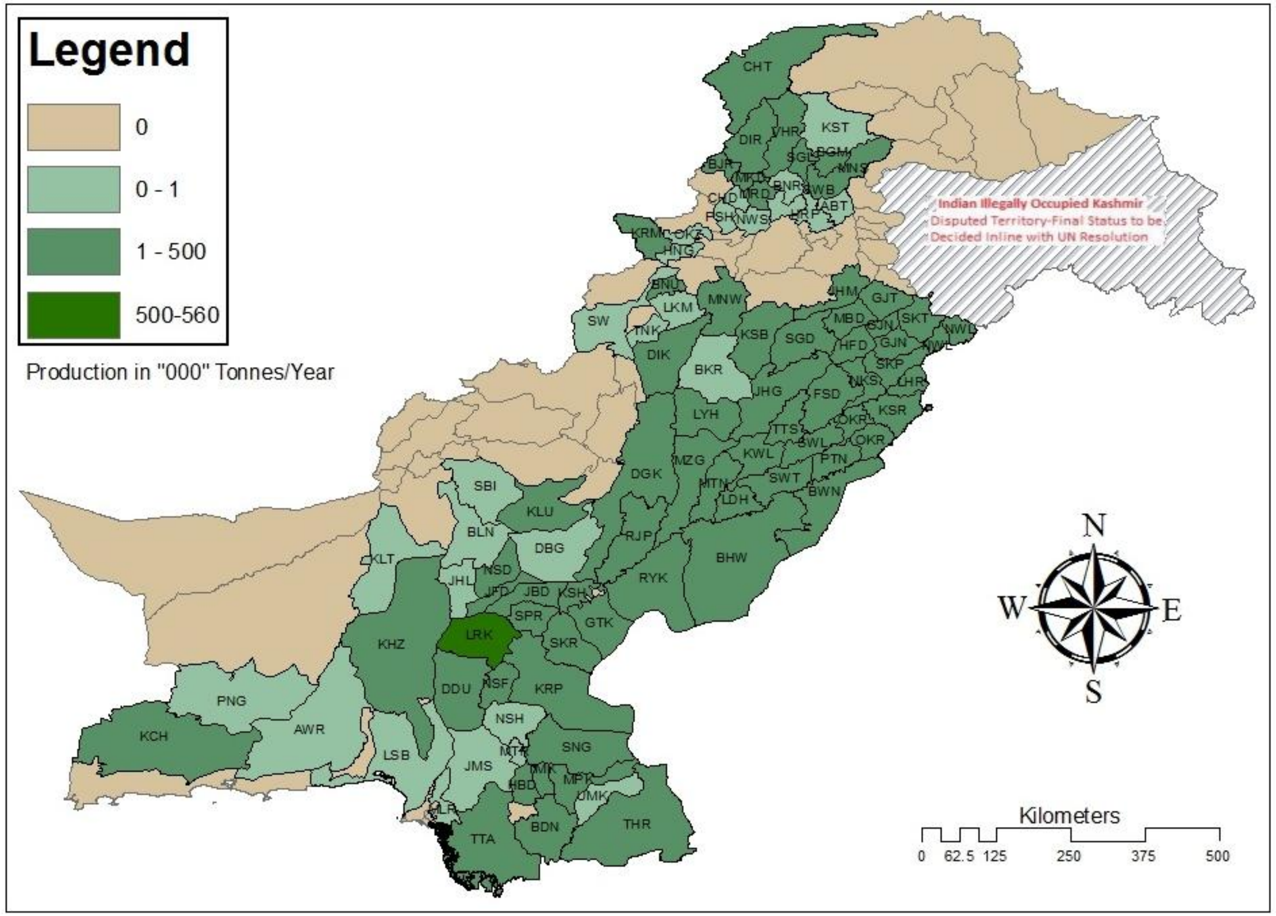
- Bulk of wheat production (500-770 thousand tonnes) is done in Rahimyar Khan, Bhawalpur, Bhawalnagar, Vehari, Muzaffargarh, Jhang, Faisalabad, Okara and Sahiwal districts.
- Rest of districts' production fall in the range of 1-500 thousand tonnes/year.
- Extreme northeast of the country where no wheat crop is grown.




(Data source: PBS)

# District-wise *Rice* production in Pakistan

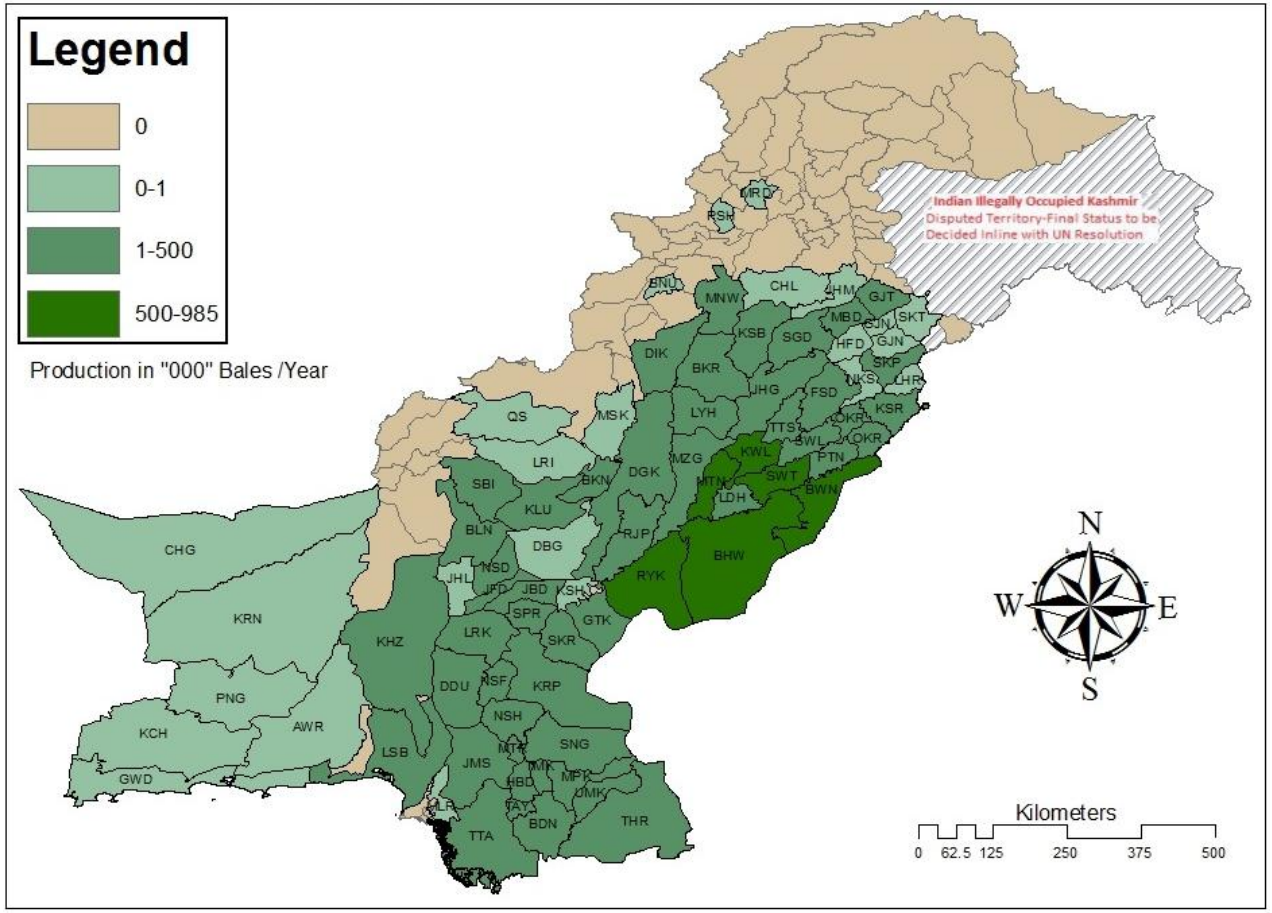
- Rice production is maximum (over 500 thousand tonnes per year) in Larkana district
- Most of the rest districts in Punjab, KP and Sindh have annual production =<500 thousand tonnes and
- Some  districts in KP, Western Sindh and Balochistan produce 1000 tonnes/year.




(Data source: PBS)

# District-wise Cotton production across Pakistan

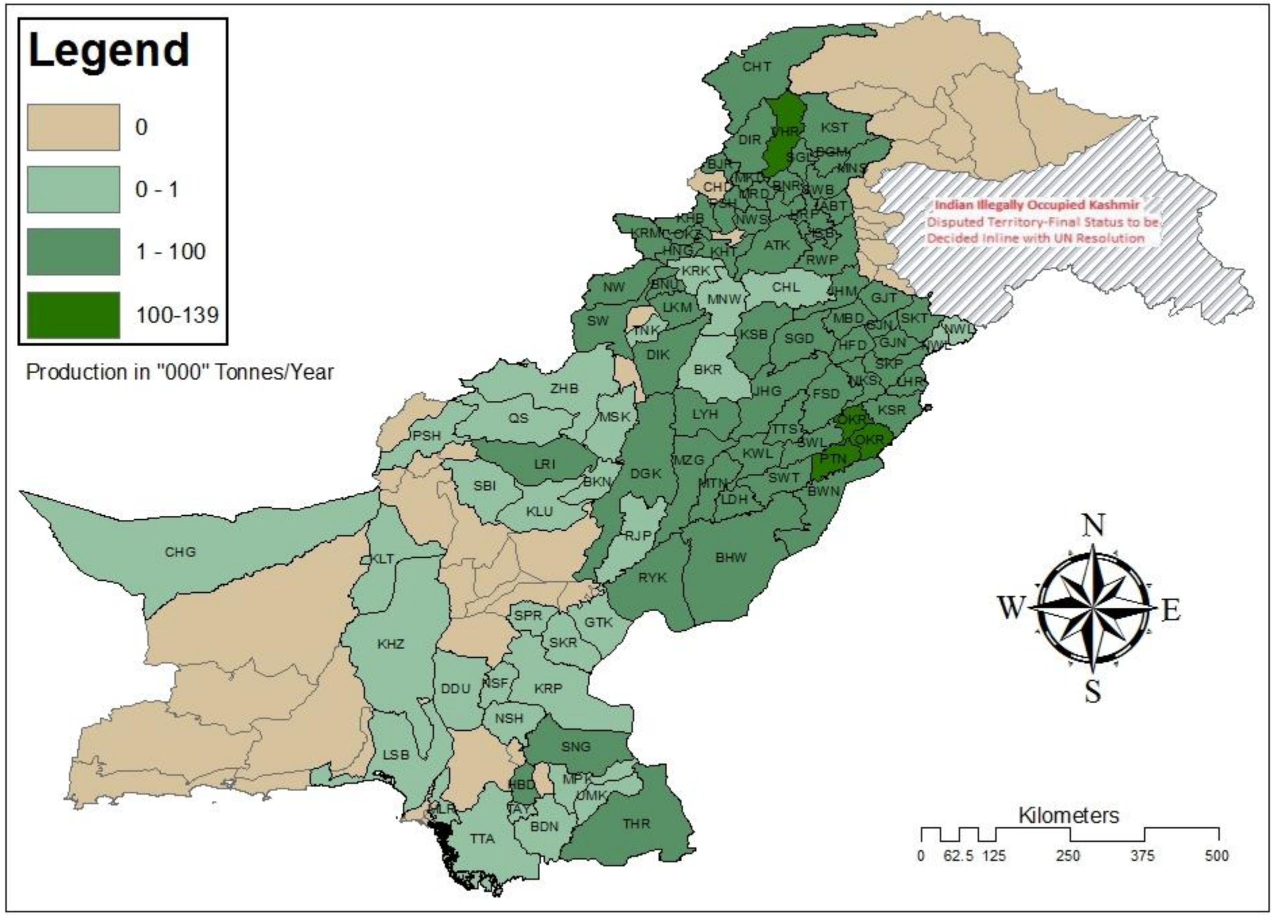
- Maximum of =>500-985 thousands bales per year is produced in Bhawalpur, Bhawalnagar, Multan, Khaniwal and Vehari districts.
- Most districts in central & south Punjab, Sindh and east/northeast Balochistan have a growth of 1000-500,000 tonnes bales per year.




(Data source: PBS)

# District-wise *Maize* production across Pakistan

- Three districts in Centre-East and one in Northwest produce maximum Maize crop of 100,000-139,000 tonnes/year.
- Most districts Punjab, KP and Southeast Sindh producing 100 thousand tonnes annually.




(Data source: PBS)

# Wheat-climate best-fit regression models and their rankings based on AIC

| Dependent variable (predictand) | Model Notation | Model Description (predictors) | df | AIC | AICc | Δi | Model Ranking |
|---|---|---|---|---|---|---|---|
| Wheat yield | $M_1$ | $T_{max}+T_m$ | 2 | 831.3 | 828.5 | 0.0 | 1 |
| | $M_2$ | $T_{max} + T_m + R_1$ | 3 | 832.1 | 828.0 | 0.1 | 2 |
| | $M_3$ | $T_{min} + T_{max} + T_m + R_2$ | 4 | 832.9 | 827.8 | 0.2 | 3 |
| | $M_4$ | $T_{min} + T_{max} + T_m + R_1$ | 4 | 833.4 | 829.2 | 0.7 | 4 |
| | $M_5$ | $T_{min}+T_m+R_1$ | 3 | 832.9 | 829.4 | 0.9 | 5 |
| | $M_6$ | $T_{min}+T_{max}+T_m$ | 3 | 833.2 | 829.7 | 1.2 | 6 |
| | $M_7$ | $T_m+R_1$ | 2 | 838.8 | 836.0 | 7.5 | 7 |

- Model $M_1$ (with Δi=0) comprising seasonal mean and seasonal maximum temperatures, is the best-fit model
- $M_2$ (Δi=0.1) with same regressors but with an addition of rainfall (of DJFM) is the 2nd best-fit model.
- Models with values Δi< 2 are considered as good as the best model.
- Models with Δi= 6 ought not be disregarded
- Models with values Δi<10 can be regarded adequate.
- So, all 7 models given here are quite adequate models.

# Rice-climate best-fit regression models and their ranking based on AIC

| Dependent variable (predictand) | Model Notation | Model Description (predictors) | df | AIC | AICc | Δi | Model Ranking |
|---|---|---|---|---|---|---|---|
| Rice yield | $M_1$ | $T_m+T_{max}+R$ | 3 | 400.1 | 401.0 | 0 | 1 |
| | $M_2$ | $T_m+T_{min}+T_{max}+R$ | 4 | 401.8 | 403.5 | 2.4 | 2 |
| | $M_3$ | $T_{min}+T_{max}+R$ | 3 | 403.4 | 404.4 | 3.3 | 3 |
| | $M_4$ | $T_{min}+R$ | 2 | 404.4 | 404.8 | 3.8 | 4 |
| | $M_5$ | $T_m+T_{max}$ | 2 | 405.0 | 405.5 | 4.5 | 5 |
| | $M_6$ | $T_m+R$ | 2 | 405.4 | 405.8 | 4.8 | 6 |
| | $M_7$ | $T_m+T_{min}+T_{max}$ | 3 | 405.6 | 406.5 | 5.5 | 7 |
| | $M_8$ | $T_{min}$ | 1 | 406.5 | 406.6 | 5.6 | 8 |
| | $M_9$ | $T_m$ | 1 | 406.5 | 406.7 | 5.6 | 8 |
| | $M_{10}$ | $T_{max}$ | 1 | 406.7 | 406.9 | 5.8 | 9 |
| | $M_{11}$ | $T_{max}+R$ | 2 | 406.5 | 407.0 | 5.9 | 10 |
| | $M_{12}$ | $T_m+T_{min}+R$ | 3 | 406.0 | 406.9 | 5.9 | 10 |
| | $M_{13}$ | $T_{min}+T_{max}$ | 2 | 406.8 | 407.2 | 6.2 | 11 |
| | $M_{14}$ | $T_m+T_{min}$ | 2 | 408.5 | 409.0 | 8.0 | 12 |

# Maize-climate best-fit regression models and their rankings based on AIC

| Dependent variable (predictand) | Model Notation | Model Description (predictors) | df | AIC | AICc | Δi | Model Ranking |
|---|---|---|---|---|---|---|---|
| Maize yield | $M_1$ | $T_m+T_{max}$ | 2 | 402.7 | 403.2 | 0 | 1 |
| | $M_2$ | $T_{min}$ | 1 | 403.8 | 404.0 | 0.8 | 2 |
| | $M_3$ | $T_m$ | 1 | 404.0 | 404.1 | 1.0 | 3 |
| | $M_4$ | $T_{max}$ | 1 | 404.2 | 404.4 | 1.2 | 4 |
| | $M_5$ | $T_{min}+T_{max}$ | 2 | 404.3 | 404.7 | 1.5 | 5 |
| | $M_6$ | $T_{min}+R$ | 2 | 404.6 | 405.1 | 1.9 | 6 |
| | $M_7$ | $T_{min}+T_m+T_{max}$ | 3 | 404.1 | 405.1 | 1.9 | 6 |
| | $M_8$ | $T_m+R$ | 2 | 405.3 | 405.7 | 2.6 | 7 |
| | $M_9$ | $T_{max}+T_{min}+R$ | 3 | 405.1 | 406.1 | 2.9 | 8 |
| | $M_{10}$ | $T_m+T_{min}$ | 2 | 405.8 | 406.3 | 3.1 | 9 |
| | $M_{11}$ | $T_{max}+R$ | 2 | 406.2 | 406.7 | 3.5 | 10 |
| | $M_{12}$ | $T_m+T_{min}+T_{max}+R$ | 4 | 405.1 | 406.7 | 3.6 | 11 |
| | $M_{13}$ | $T_m+T_{min}+R$ | 3 | 406.5 | 407.4 | 4.2 | 12 |

# Cotton-climate best-fit regression models and their rankings based on AIC

| Dependent variable (predictand) | Model Notation | Model Description (predictors) | df | AIC | AICc | Δi | Model Ranking |
|---|---|---|---|---|---|---|---|
| Cotton yield | $M_1$ | $T_{min}$ | 1 | 540.6 | 540.7 | 0 | 1 |
| | $M_2$ | $T_m$ | 1 | 540.8 | 540.9 | 0.2 | 2 |
| | $M_3$ | $T_{max}$ | 1 | 540.8 | 540.9 | 0.2 | 2 |
| | $M_4$ | $T_{min}+T_m+T_{max}$ | 3 | 540.2 | 541.0 | 0.3 | 3 |
| | $M_5$ | $T_m$ +R | 2 | 542.3 | 542.7 | 2.0 | 4 |
| | $M_6$ | $T_{min}$+R | 2 | 542.4 | 542.8 | 2.0 | 4 |
| | $M_7$ | $T_{max}$+R | 2 | 542.6 | 543.0 | 2.3 | 5 |
| | $M_8$ | $T_{min}+T_{max}$ | 2 | 542.8 | 543.2 | 2.4 | 6 |
| | $M_9$ | $T_m+T_{max}$ | 2 | 542.8 | 543.2 | 2.5 | 7 |
| | $M_{10}$ | $T_{max}+T_{min}$+R | 3 | 544.3 | 545.1 | 4.3 | 8 |
| | $M_{11}$ | $T_m+T_{max}$+R | 3 | 544.3 | 545.1 | 4.4 | 8 |
| | $M_{12}$ | $T_m+T_{min}$+R | 3 | 544.4 | 545.2 | 4.4 | 9 |

# Seasonal (JAS, DJFM & AMJ) Rainfalls Trends

| No | Station | S Statistic | Z value | p-value | Test Result |
|---|---|---|---|---|---|
| | **a. Monsoon (JAS) rainfall trend** | | | | |
| 1 | Bunji | 355 | 2.57 | 0.01** | Statistically significant increasing trend |
| 2 | Chilas | 448 | 3.2453 | 0.001** | Do |
| 3 | Gupis | 271 | 1.9601 | 0.04* | Do |
| 4 | Skardu | 421 | 3.0494 | 0.002** | Do |
| 5 | Islamabad | 332 | 2.403 | 0.02* | Do |
| 6 | Risalpur | 315 | 2.2795 | 0.02* | Do |
| 7 | Peshawar | 319 | 2.3087 | 0.02* | Do |
| 8 | Mianwali | 453 | 3.2813 | 0.001** | Do |
| 9 | Dalbandin | -249 | 2.0134 | 0.04* | Statistically significant decreasing trend |
| 10 | Garhi-dupatta | -397 | 2.8784 | 0.004** | Do |
| 11 | Jiwani | -255 | 2.0622 | 0.04* | Do |
| | **b. Winter (DJFM) rainfall trend** | | | | |
| 1 | Dir | 301 | 2.1779 | 0.03* | Statistically significant increasing trend |
| 2 | Kalat | 291 | 2.1057 | 0.04* | Do |
| 3 | Mianwali | 287 | 2.1338 | 0.03* | Do |

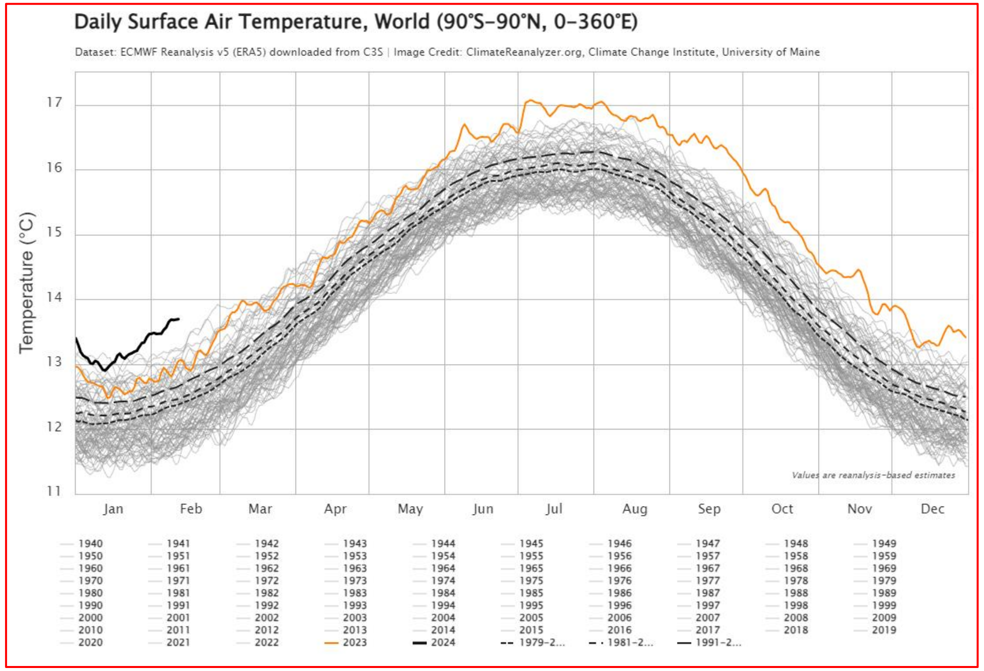
| Station | S Statistic | Z value | p-value | Test Result |
|---|---|---|---|---|
| **c. Spring (AMJ) rainfall trend** | | | | |
| Barkhan | 454 | 3.2888 | 0.001** | Statistically significant increasing |
| Bhawalnagar | 463 | 3.3541 | 0.0008** | Do |
| Bahawalpur | 276 | 1.9973 | 0.05* | Do |
| Khuzdar | 278 | 2.0109 | 0.04* | Do |
| Lahore | 284 | 2.0545 | 0.04* | Do |
| Lasbella | 458 | 3.3207 | 0.0008** | Do |
| Mianwali | 348 | 2.5191 | 0.01** | Do |
| Chitral | -335 | 2.4247 | 0.02* | Statistically significant decreasing trend |
| Drosh | -380 | 2.7514 | 0.006** | Do |
| Garhi-dupatta | -407 | 2.9474 | 0.003** | Do |

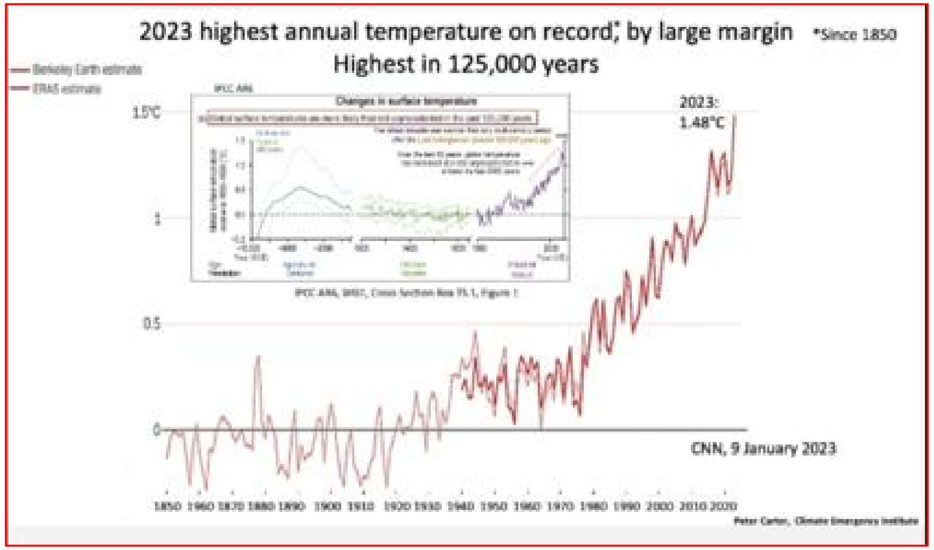
** Significant at 99% confidence interval and * significant at 95% confidence interval

# Future climate scenarios

(Source: Sarfaraz, 2021, PhD Dissertation)

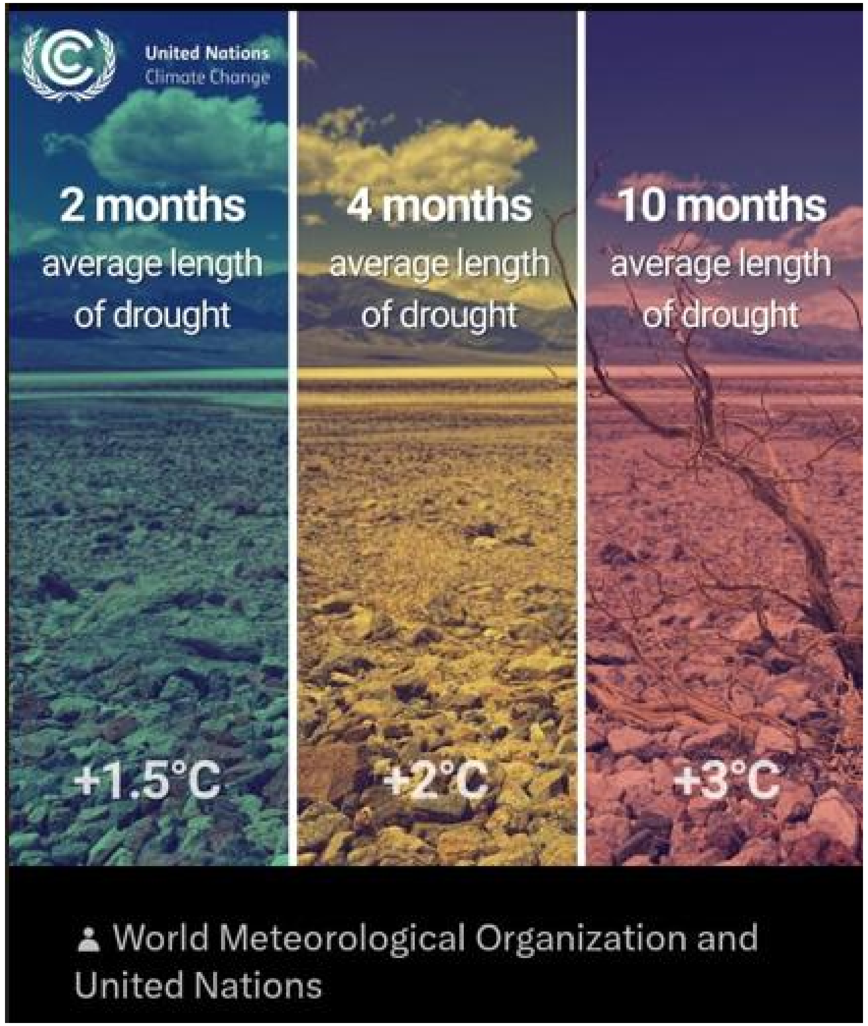
# Global warming and future drought




UN/WMO estimate that

- At 1.5°C global temperature increase, the world can experience 2 months average length of drought
- At 2°C, it could be 4 months and
- At 3°C, it could be 10 months average length of drought.



# Climate Risks: 1.5°C vs 2°C Global Warming

| Climate Risk | 1.5°C | 2°C |
|---|---|---|
| Extreme Weather | **100%** increase in flood risk | **170%** increase in flood risk |
| Water Availability | **350 million** urban residents exposed to severe drought by 2100 | **410 million** urban residents exposed to severe drought by 2100 |
| Arctic Sea Ice | Ice-free summers in the Arctic at least once **every 100 years** | Ice-free summers in the Arctic at least once **every 10 years** |
| People | **9%** of the world's population (**700m people**)will be exposed to extreme heatwaves at least once every **20 years** | **28%** of the world's population (**2 billion people**)will be exposed to extreme heatwaves at least once every **20 years** |
| Sea-level Rise | **46 million people** impacted by sea-level rise of **48cm** by 2100 | **49 million people** impacted by sea-level rise of **56cm** by 2100 |
| Oceans | **Lower risks** to marine biodiversity ecosystems and their ecological functions & services | **Higher risks** to marine biodiversity ecosystems and their ecological functions & services |
| Coral Bleaching | **70%** of the **coral reefs** are lost by 2100 | **Virtually all the coral reefs are lost** by 2100 |
| Costs | **Lower economic growth at 2°C** than **at 1.5°C** for many countries, particularly for low-income countries | |
| Food | Every half degree warming will consistently lead to lower yields and lower nutritional content in tropical regions | |

# Climate change impacts

(Source: Janjua et al. 2014, Tariq et al. 2010, Shakoor et al. 2011)

# Conclusion

- For Wheat yiled, the AIC-based regression modeling indicates that seasonal Tm + Tmax are the best regressors, followed by the 6 other models with different seasonal climate indices.
- For Rice yield, the AIC-based regression modeling indicates that the seasonal $T_m$+$T_{max}$+R is the best-fit model followed by 13 other models with different seasonal climate indices combinations.
- For Maize yield, the AIC-based regression modeling indicates that the seasonal Tm + Tmax are the best-fit model regressors followed by 12 other models with different seasonal climate indices combinations.
- For Cotton yield, the AIC-based regression modeling indicates that seasonal Tmin is found in the best-fit model followed by 11 rest models with different seasonal climate indices combinations.
- Seasonal mean temperature (Tm) with significant ($p=0.015$) increasing trend may reduce the wheat yield which can jeopardize the national food security and seasonal rainfall (R1) potentially erratic may also affect the wheat as well as other crops yield production.
- At 1.5°C global temperature increase, the world can experience 2 months average length of drought
- At 2°C, it could be 4 months and
- At 3°C, it could be 10 months average length of drought.

# References:

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. In: (Eds.): Petrov, B.N. and F. Caski Proc. Second International Symposium on Information Theory. Akademiai Kiado. Budapest, 267-281.

Akaike, H. 1983b. Information measures and model selection. Int. Statistical Institute. 44: 277-291. Akaike, H. 1992. Information theory and an extension of the maximum likelihood principle. In: (Eds.): Kotz, S. and N.L. Johnson, Breakthroughs in Statistics., 1: 610-624.

Akaike, H. 1994. Implications of the informational point of view on the development of statistical science. In: (Ed.): Bozdogan, H. Engineering and scientific applications. Proc. First US/Japan conference on the frontiers of statistical modeling: an informational approach. Kluwer, Dordrecht., 3: 27-38.

Burnham, K.P. and D.R. Anderson. 2001. Kullback-Leibler information as a basis for strong inference in ecological studies. Wildl. Res., 28: 111-119.

Burnham, K.P. and D.R. Anderson. 2002. Model Selection and Multimodel Inference. Springer, New York (Print).

Hurvich, C. and C.L. Tsai. 1989. Regression and time series model selection in small samples. Biometrika., 76: 297-307.

Janjua, P.Z., G. Samad and N. Khan. 2014. Climate change and wheat production in Pakistan: An autoregressive distributed lag approach. NJAS-Wageningen J. Life Sci., 68: 13-19.

Sarfaraz, S., Shoukat, S. S., & Khan, T. M. A. 2021. Investigating wheat yield and climate parameters regression model based on Akaike information criteria. *Pak. J. Bot*, *53*(4), 1299-1306.

Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions. J. R. Stat. Soc., l: 111-147.

Stone, M. 1977. An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. J. R. Stat. Soc., 1: 44-47.

Shakoor, U., A. Saboor, I. Ali and A.Q. Mohsin. 2011. Impact of climate change on agriculture: empirical evidence from arid region. Pak. J. Agri. Sci., 48: 327-333.

Tariq, A., N. Tabasam, K. Bakhsh, M. Ashfaq and S. Hassan. 2014. Food security in the context of climate change in Pakistan. PJCSS., 8: 540-550.

# *Thank You !*