"O GLM, assim como o GLMM não possui a premissa de linearidade entre as variáveis?"

A resposta é: **Sim, eles possuem uma premissa de linearidade, mas ela é mais flexível** do que na regressão linear simples (lm).

* **Regressão Linear Simples (lm)**: Assume que a **relação entre a variável resposta (Y) e a explicativa (X) é uma linha reta**. Y = β₀ + β₁X.
* **GLM / GLMM**: Não assume que a relação entre Y e X é uma linha reta. Em vez disso, assume que a relação entre os preditores (X) e uma **transformação da média da resposta (através da função de ligação)** é linear.

**Exemplo Prático (seu caso):** Em um GLM de Poisson ou Tweedie com uma função de ligação logarítmica (o padrão), o modelo assume que: log(Média de Y) = β₀ + β₁X

Isso significa que o modelo espera uma **relação exponencial** entre Y e X na escala original dos dados, não uma linha reta. Se você plotar Y vs. X, verá uma curva. Mas se você plotar log(Y) vs. X, a relação deveria ser aproximadamente linear.

Portanto, a premissa do GLM/GLMM não é tão rígida, mas ainda assume uma forma monotônica e específica (exponencial, logística, etc.) para a relação.

**Quando um GAMM é a Melhor Escolha?**

Você observou que sua variável resposta não possui uma relação linear com as explicativas. Se, além disso, a relação também **não segue uma curva simples e monotônica** (como a exponencial do modelo Poisson/Tweedie), então um GLMM pode não ser suficiente.

É aqui que o **GAMM (Modelo Aditivo Generalizado Misto)** brilha.

Um GAMM é uma extensão de um GLMM. Em vez de forçar a relação a seguir uma forma linear na escala da função de ligação (ex: β₁ \* Tmean), ele substitui esse termo por uma **função de suavização flexível** (ex: s(Tmean)).

**Analogia:**

* **GLMM:** Tenta ajustar uma "régua reta" à relação entre log(Y) e X.
* **GAMM:** Usa uma "régua flexível" que pode se curvar para capturar a verdadeira forma da relação, seja ela qual for (com picos, vales, platôs, etc.).

Como você disse, se a relação não é linear de forma alguma, **modelar com um GAMM é sim a abordagem mais interessante, flexível e honesta**, pois permite que os próprios dados determinem a forma da relação, em vez de forçá-la a um padrão pré-definido.

**Regras de Ouro e Cuidados Essenciais ao Usar o AIC**

Para que a comparação seja válida, você **precisa** seguir algumas regras cruciais:

1. **Os Mesmos Dados:** Todos os modelos candidatos devem ser ajustados exatamente ao **mesmo conjunto de dados**. Se houver valores ausentes (NA), certifique-se de que o número de observações seja idêntico para todos os modelos.
2. **Mesma Variável Resposta e Família:** Você não pode usar o AIC para comparar um modelo para log(eggs.fem.day) com um modelo para eggs.fem.day. A variável resposta e a família de distribuição (ex: Tweedie, Binomial Negativa) devem ser as mesmas em todos os modelos.
3. **Cuidado com ML vs. REML (CRÍTICO para Modelos Mistos):**
   * Modelos mistos (GLMM/GAMM) podem ser ajustados por dois métodos: REML (Máxima Verossimilhança Restrita) e ML (Máxima Verossimilhança).
   * **Regra:** Se você está comparando modelos com **diferentes estruturas de efeitos fixos** (que é o seu caso, ao adicionar ou remover preditores e interações), você **DEVE** ajustar todos os modelos usando **Máxima Verossimilhança (ML)**.
   * O padrão em muitas funções (como lme4::lmer ou mgcv::gam) é REML. Você precisa especificar manualmente para usar ML.

 O método REML (Máxima Verossimilhança Restrita) é excelente para estimar os componentes de variância (os efeitos aleatórios e os parâmetros de suavização) de forma menos enviesada. No entanto, ele faz isso de uma forma que torna a comparação da parte dos efeitos fixos entre modelos diferentes inválida. A Máxima Verossimilhança (ML) permite uma comparação justa do modelo como um todo.

 **Boas Práticas:**

1. **Para seleção de modelos:** Ajuste todos os seus modelos candidatos com REML = FALSE (ou method = 'ML') para compará-los com o AIC.
2. **Para o relatório final:** Depois de escolher a melhor estrutura de efeitos fixos, muitos estatísticos recomendam reajustar esse modelo final usando REML = TRUE (o padrão) para obter as estimativas finais dos parâmetros, pois o REML é melhor para estimar os componentes de variância.